

ANALYSE GÉNÉTIQUE MODERNE

• GRIFFITHS • GELBART • MILLER • LEWONTIN •

Traduction de la 1^{re} édition américaine par Chrystelle Sanlaville
Révision scientifique de Dominique Charmot et Denise Aragnol



DeBoeck 



Bibliothèque Scientifique

Groupe « ouvert »

facebook

ANALYSE GÉNÉTIQUE MODERNE



Bibliothèque Scientifique

Groupe « ouvert »

facebook

**Chez le même éditeur
Extrait du catalogue**

Génétique

LEWIN B., *Gènes VI*

PLOMIN R., DEFRIES J.C., MCCLEARN G.E., RUTTER M., *Des gènes au comportement. Introduction à la génétique comportementale*

WATSON J.D., GILMAN M., WITKOWSKI J., ZOLLER M., *ADN recombinant*

ANALYSE GÉNÉTIQUE MODERNE

- Anthony J.F. GRIFFITHS
University of British Columbia
- Williams M. GELBART
Harvard University
- Jeffrey H. MILLER
University of California, Los Angeles
- Richard C. LEWONTIN
Harvard University

Traduction de la 1^{re} édition américaine par Chrystelle Sanlaville
Révision scientifique de Dominique Charmot (Université Aix Marseille)
et Denise Aragnol (Université Aix Marseille)



Ouvrage original :

Modern Genetic Analysis by Anthony J.F. Griffiths, William M. Gelbart, Jeffrey H. Miller, Richard C. Lewontin

First published in the United States by W.H. Freeman and Company, New York, New York and Basingstoke

Publié pour la première fois aux États-Unis par W.H. Freeman and Company, New York, New York and Basingstoke

© 1999 by W.H. Freeman and Company

All rights reserved.

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : <http://www.deboeck.be>

© De Boeck Université s.a., 2001
171, rue de Rennes, F-75006 Paris
rue des Minimes 39, B-1000 Bruxelles
Pour la traduction et l'adaptation française

1^{re} édition

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé en Italie

Dépôt légal :
Bibliothèque Nationale, Paris : mars 2001
Bibliothèque Royale Albert 1er, Bruxelles : 2001/0074/73

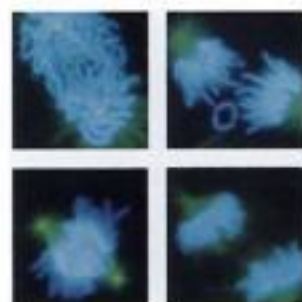
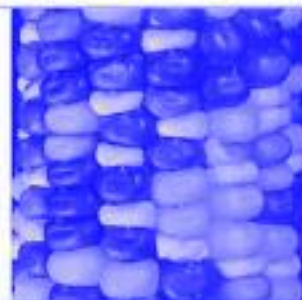
ISBN : 2-7445-0111-5

Sommaire

<u>1</u>	<u>La génétique et l'organisme</u>	<u>1</u>
<u>2</u>	<u>La structure des gènes et des génomes</u>	<u>23</u>
<u>3</u>	<u>La fonction des gènes</u>	<u>51</u>
<u>4</u>	<u>La transmission des gènes</u>	<u>85</u>
<u>5</u>	<u>La recombinaison des gènes</u>	<u>129</u>
<u>6</u>	<u>L'interaction des gènes</u>	<u>165</u>
<u>7</u>	<u>Les mutations géniques</u>	<u>197</u>
<u>8</u>	<u>Les mutations chromosomiques</u>	<u>235</u>
<u>9</u>	<u>La génétique des bactéries et des phages</u>	<u>271</u>
<u>10</u>	<u>La technologie de l'ADN recombinant</u>	<u>299</u>
<u>11</u>	<u>Les applications de la technologie de l'ADN recombinant</u>	<u>341</u>
<u>12</u>	<u>La génomique</u>	<u>373</u>
<u>13</u>	<u>Les éléments génétiques transposables</u>	<u>413</u>
<u>14</u>	<u>La régulation de la transcription des gènes</u>	<u>433</u>
<u>15</u>	<u>La régulation du nombre de cellules :</u>	
	<u>les cellules normales et les cellules cancéreuses</u>	<u>465</u>
<u>16</u>	<u>Les fondements génétiques du développement</u>	<u>491</u>
<u>17</u>	<u>La génétique des populations et l'évolution</u>	<u>535</u>
<u>18</u>	<u>La génétique quantitative</u>	<u>575</u>
	Glossaire	607
	Lectures conseillées	635
	Réponses à quelques problèmes	645
	Index	659

Table des matières

1	LA GÉNÉTIQUE ET L'ORGANISME	1
	La génétique et les activités humaines	3
	La génétique et la biologie	10
	La variation est à la base de la génétique	14
	La génétique expérimentale	17
2	LA STRUCTURE DES GÈNES ET DES GÉNOMES	23
	La nature de l'ADN	24
	La nature des gènes	31
	La nature des génomes	33
	La nature des chromosomes nucléaires eucaryotes	39
3	LA FONCTION DES GÈNES	51
	Les gènes et les ARN	52
	Fabriquer des transcrits fonctionnels	54
	La structure d'une protéine	61
	La traduction	64
	Fonctionnement et dysfonctionnement des protéines dans les cellules	68
	Protéines déficientes, dominance et récessivité	77
	La division fonctionnelle des tâches dans l'ensemble des gènes	78
4	LA TRANSMISSION DES GÈNES	85
	La réplication de l'ADN	86
	La division cellulaire	91
	Les modes de transmission des gènes individuels	101



L'analyse des arbres généalogiques humains	111
La transmission des gènes d'organites	118
5 LA RECOMBINAISON DES GÈNES	129
L'assortiment indépendant	131
Le crossing-over	137
Les cartes de liaison génétique	140
Le crossing-over mitotique	151
Le mécanisme du crossing-over	152
La recombinaison à l'intérieur d'un gène	153
6 L'INTERACTION DES GÈNES	165
Des gènes aux phénotypes	166
Un test de diagnostic pour les allèles	166
Les interactions entre les allèles d'un même gène	170
L'interaction des gènes conduit à des proportions modifiées de dihybrides	174
La pénétrance et l'expressivité	182
7 LES MUTATIONS GÉNIQUES	197
Les fondements moléculaires de la mutation	198
L'analyse des mutations	221
8 LES MUTATIONS CHROMOSOMIQUES	235
Les changements du nombre de chromosomes	236
Les réarrangements chromosomiques	248
Les conséquences globales des mutations chromosomiques chez l'homme	258
L'évolution du génome	258
9 LA GÉNÉTIQUE DES BACTÉRIES ET DES PHAGES	271
Travailler avec des micro-organismes	272
La conjugaison bactérienne	274
La transformation bactérienne	282
La génétique des bactériophages	283
La transduction	288
Un résumé du transfert des gènes bactériens	290



10	LA TECHNOLOGIE DE L'ADN RECOMBINANT	299
	La construction de l'ADN recombinant	300
	Cloner un gène spécifique	307
	Utiliser de l'ADN cloné	320
11	LES APPLICATIONS DE LA TECHNOLOGIE DE L'ADN RECOMBINANT	341
	La mutagenèse <i>in vitro</i>	342
	La cartographie à l'aide des RFLP	342
	La génétique inverse	346
	Exprimer des gènes eucaryotes dans des bactéries	346
	La technologie de l'ADN recombinant chez les Eucaryotes	349
	La thérapie génique	361
	Utiliser l'ADN recombinant pour détecter directement les allèles responsables de maladies	365
12	LA GÉNOMIQUE	373
	La génomique structurale	376
	La génomique fonctionnelle	399
13	LES ÉLÉMENTS GÉNÉTIQUES TRANSPOSABLES	413
	Les séquences d'insertion	414
	Les transposons	417
	Le mécanisme de la transposition	419
	Les réarrangements par l'intermédiaire des éléments transposables	421
	Un résumé des éléments génétiques transposables chez les Procaryotes	422
	Les éléments Ty de la levure	422
	Les éléments transposables de la drosophile	423
	Les rétrovirus	425
	La transposition par l'intermédiaire d'un ARN	426
	Les éléments de contrôle chez le maïs	427
	Un résumé des propriétés des éléments transposables chez les Eucaryotes	431
14	LA RÉGULATION DE LA TRANSCRIPTION DES GÈNES	433
	La logique de la régulation des gènes procaryotes	434



<u>L'essentiel de la régulation de la transcription</u>	
<u>chez les Procaryotes</u>	434
<u>La régulation du système lactose</u>	436
<u>Le double contrôle positif et négatif : l'opéron arabinose</u>	446
<u>Les voies métaboliques</u>	447
<u>La transcription : une vue d'ensemble de la régulation</u>	
<u>des gènes chez les Eucaryotes</u>	448
<u>La régulation des facteurs de transcription</u>	456
<u>La transmission épigénétique</u>	459

15 LA RÉGULATION DU NOMBRE DE CELLULES : LES CELLULES NORMALES ET LES CELLULES CANCÉREUSES

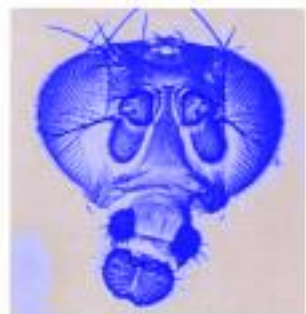
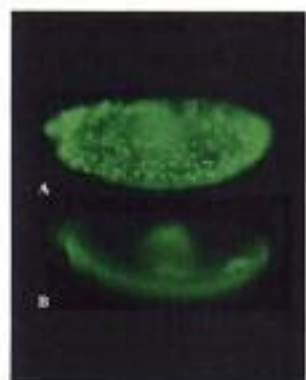
<u>Le cancer et le contrôle du nombre de cellules : une vue d'ensemble</u>	466
<u>La machinerie responsable de la prolifération cellulaire</u>	467
<u>La machinerie de la mort cellulaire programmée</u>	470
<u>Contrôler la prolifération cellulaire et la machinerie de l'apoptose</u>	472
<u>Le cancer : la génétique du contrôle cellulaire aberrant</u>	478

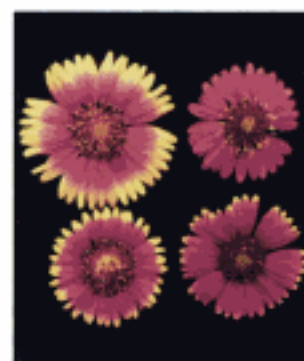
16 LES FONDEMENTS GÉNÉTIQUES DU DÉVELOPPEMENT

<u>La problématique de la génétique du développement</u>	492
<u>Les choix binaires : les voies de la détermination du sexe</u>	495
<u>La détermination du sexe chez la drosophile : chaque cellule décide seule</u>	496
<u>La détermination du sexe chez les mammifères : le contrôle coordonné par le système endocrine</u>	501
<u>Les choix binaires : lignée germinale ou lignée somatique</u>	504
<u>Établir un plan complexe de construction : la mise en place de l'information de position</u>	508
<u>La formation d'un plan complexe : utiliser l'information de position pour établir des identités cellulaires</u>	515
<u>D'autres aspects de la formation du plan d'organisation</u>	523
<u>Les nombreux parallèles entre la formation des plans d'organisation chez les vertébrés et les insectes</u>	527

17 LA GÉNÉTIQUE DES POPULATIONS ET L'ÉVOLUTION

<u>La révolution de Darwin</u>	536
<u>La variation et sa modulation</u>	537





Les conséquences de la reproduction sexuée sur la variation	544
Les origines de la variation	546
La sélection	553
Le polymorphisme équilibré	558
Les sommets adaptatifs multiples	560
La sélection artificielle	562
Les événements aléatoires	563
Une synthèse des forces évolutives	565
 18 LA GÉNÉTIQUE QUANTITATIVE	 575
Quelques notions de statistique	577
Les génotypes et la distribution phénotypique	579
La norme de réaction et la distribution phénotypique	581
Déterminer les normes de réaction	582
L'héritabilité d'un caractère	584
Quantifier l'héritabilité	586
Localiser les gènes	590
Une analyse approfondie de la variance	592
Appendice statistique au chapitre 18	596
 GLOSSAIRE	 607
 LECTURES CONSEILLÉES	 635
 RÉPONSES À QUELQUES PROBLÈMES	 645
 INDEX	 659

Avant-propos

EN TANT QUE PROFESSEURS DE GÉNÉTIQUE

et auteurs de *Introduction à l'analyse génétique*, nous sommes témoins de la révolution silencieuse de l'enseignement de la génétique aux étudiants de fraîche date. De nombreux enseignants trouvent qu'une approche strictement chronologique, ou historique, n'est plus adaptée à leur façon d'enseigner la génétique et ne correspond pas non plus aux besoins de leurs étudiants. La génétique moléculaire est introduite de plus en plus tôt dans les cours et intégrée à l'analyse phénotypique et génotypique.

Analyse génétique moderne a été écrit à l'attention des enseignants et des étudiants qui ont besoin d'un ouvrage étayant l'approche « L'ADN d'abord ». Se démarquer de l'approche historique traditionnelle de la génétique a eu quelques conséquences significatives sur la rédaction de l'ouvrage, parmi lesquelles et non la moindre, une présentation plus structurée dans laquelle les principes génétiques ressortent mieux.

Que la présentation soit traditionnelle ou moderne, il est essentiel que les étudiants apprennent à raisonner comme des généticiens. C'est pourquoi, comme dans *Introduction à l'analyse génétique*, l'analyse des données et les conclusions tirées de celles-ci par les étudiants sont au centre de nos préoccupations.

Une approche moderne

Après un chapitre d'introduction sur les relations qui existent entre les gènes et les phénotypes, vient un chapitre sur la structure des gènes et des génomes. Ceci permet de définir un cadre pour la nature moléculaire du matériel héréditaire sur laquelle s'appuient ses diverses propriétés. Les chapitres 3 à 5 traitent de la fonction des gènes et de leur transmission, qui sont rapportées à la nature structurale du génome. Le chapitre 4, « La transmission des gènes », présente les différents modes possibles de transmission de gènes uniques. À un niveau supérieur de complexité, le chapitre 5, « La recombinaison des gènes », expose les principes en vigueur lors de l'analyse de deux gènes ou plus. Le niveau de complexité du chapitre 6, « L'interaction des gènes », augmente encore avec la mise en relation des modes de transmission des gènes et des interactions des gènes fonctionnels au niveau moléculaire.

Ainsi, les six premiers chapitres contiennent l'essentiel de notre approche moderne de la structure et de la fonction des gènes. Toutefois, au lieu de confiner le travail de Mendel dans un chapitre isolé, nous avons dégagé les principes déduits par Mendel et nous les avons intégrés à l'emplacement adéquat du point de vue pédagogique, aux côtés des structures cellulaires et moléculaires qui les conditionnent. Nous pensons que cette nouvelle organisation des données aidera les étudiants qui ont des difficultés à replacer la



Figure 1-10

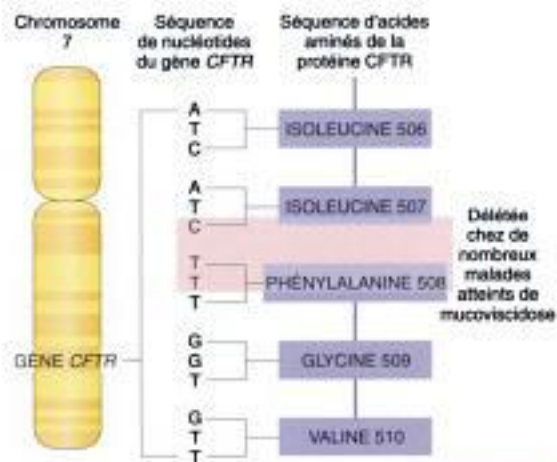


Figure 3-28

biologie moléculaire dans la charpente constituée par la génétique mendélienne. De plus, la biologie moléculaire étant introduite en premier, nous avons l'opportunité d'une intégration des concepts, mettant ainsi en relation l'organisme, la cellule, le chromosome, la protéine et l'ADN. Nous espérons enfin que cette nouvelle approche permettra aux étudiants de comprendre la façon dont la génétique est abordée dans le « monde réel », où les approches classique et moléculaire ne sont pas suivies séparément mais sont complémentaires l'une de l'autre.

En gardant à l'esprit le titre du livre, nous nous sommes efforcés de mettre à jour l'ensemble du texte. Nous avons ajouté les dernières découvertes de la génomique, dont un paragraphe sur la génomique fonctionnelle intégrant la technologie des puces d'ADN et le système de levure dit « double hybride ». La mort cellulaire programmée est également abordée dans le cadre de la biologie normale et de la génétique du cancer.

Les principes mis en valeur

L'un des buts essentiels dans l'écriture de *Analyse génétique moderne* a été que la compréhension et l'application des principes génétiques centraux soient prépondérantes sur les détails historiques. Nous espérons que les étudiants reconnaîtront et comprendront plus facilement les principes et thèmes fondamentaux s'ils ne sont pas encombrés de détails excessifs. Nous avons donc voulu réaliser une synthèse des principes de la génétique plutôt que des expériences historiques qui ont permis de les découvrir. Les principes sont utilisés pour expliquer aux étudiants comment la génétique est pratiquée aujourd'hui. Par exemple, au chapitre 3, nous introduisons les grands thèmes de la complémentarité des séquences d'acides nucléiques et de la spécificité des interactions protéines-acides nucléiques. Les étudiants verront ces sujets dans des chapitres ultérieurs, appliqués aux analyses de la réplication de l'ADN, de la synthèse protéique et de la régulation de l'expression des gènes.

Dans tous les chapitres, nous insistons sur la relation verticale qui intègre l'ADN, les protéines et les phénotypes. Les chapitres sur la technologie de l'ADN recombinant mettent l'accent sur la façon dont elle est utilisée pour isoler et caractériser des gènes plutôt que sur les techniques elles-mêmes. Les étudiants verront ainsi de quelle façon les techniques de l'ADN recombinant ont permis de cloner les gènes humains de l'albinisme et de l'alcapturie. Au chapitre 11, « Les applications de la technologie de l'ADN recombinant », un paragraphe traite des espèces transgéniques utilisées actuellement en agriculture. Le chapitre sur la génétique du développement souligne l'importance des cascades de transduction du signal dans tous les aspects du dé-

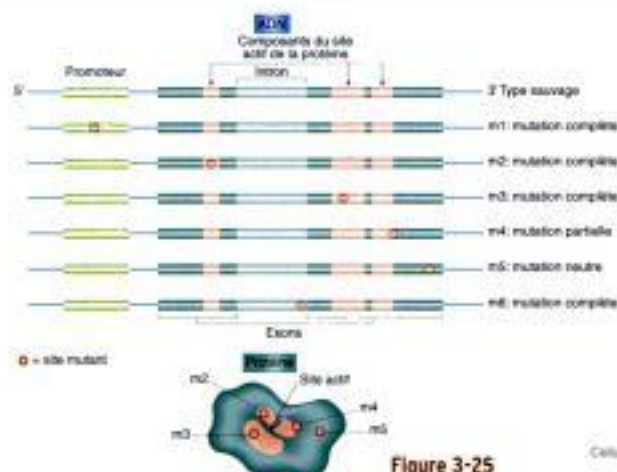


Figure 3-25

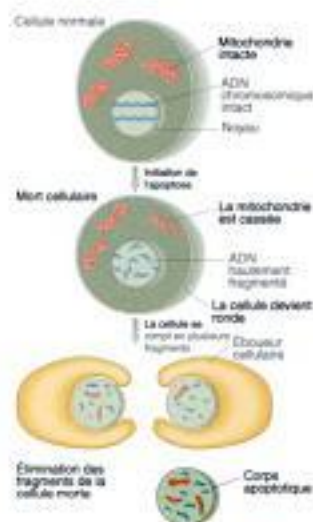


Figure 15-4

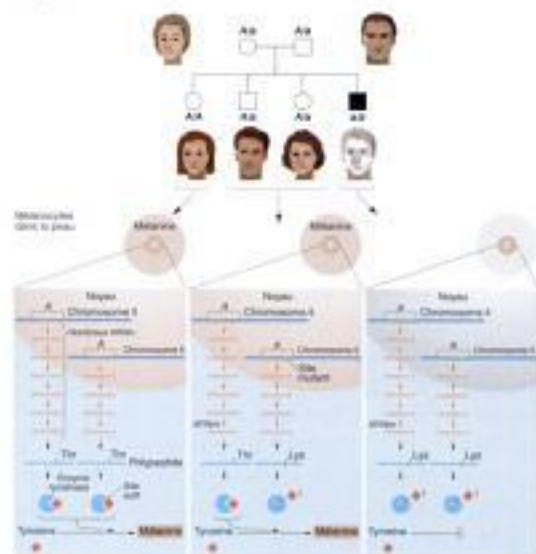


Figure 4-20

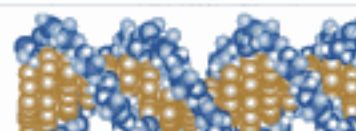
veloppement d'une cellule ou d'un organisme ainsi que les mécanismes régulateurs importants et variés, sous-jacents à toutes les décisions développementales.

Bien que nous souhaitons que les détails historiques ne détournent pas l'attention des principes essentiels, il est enrichissant pour avoir une vision complète de la génétique, que les étudiants en connaissent certaines des découvertes essentielles. Il est profitable aux étudiants d'apprendre de quelle façon ces expériences clés ont été imaginées et réalisées. Ces études fondamentales sont séparées du texte, dans des encarts appelés « Les grandes découvertes de la génétique ». Les étudiants pourront apprendre dans cette partie la façon dont Archibald Garrod a découvert la nature des erreurs congénitales du métabolisme, les recherches qui ont conduit Charles Yanofsky à déduire la colinéarité de la structure des gènes et des protéines, le modèle de Watson et Crick pour la structure de l'ADN, y compris la propre description de Watson du premier assemblage de modèle métallique, ainsi que le protocole de Luria et Delbrück qui a permis la découverte de la nature aléatoire des mutations.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 2-2

James Watson et Francis Crick proposent la structure correcte de l'ADN

Un article de 1953 écrit par James Watson et Francis Crick dans la revue *Nature* commence par les deux phrases qui ouvrent une nouvelle ère dans la biologie : « Nous aimerions suggérer une structure pour le sel d'acide désoxyribonucléique (ADN). Cette structure est basée sur des caractéristiques nouvelles qui pe-



sent une maquette d'ADN en métal. Il nous décrit plus tard, en ces termes l'excitation ressentie en assemblant le modèle :

Je descendis pour voir si l'atelier pouvait aller plus vite et produire les bases puriques et pyrimidiques à la fin de l'après-midi. Seul un petit encouragement fut nécessaire pour que la tâche soit accomplie dans les heures qui suivirent. Les pièces

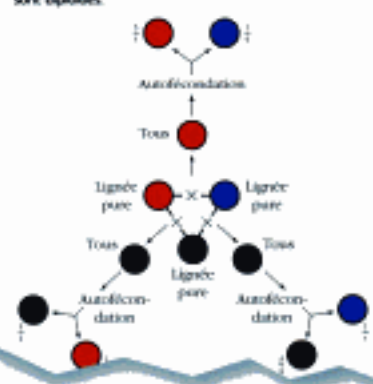
L'analyse génétique au premier plan

Nous abordons les questions qui sous-tendent la majeure partie de la génétique moderne, telles que, « Combien de gènes affectent une différence phénotypique spécifique ? Des gènes donnés sont-ils liés ? Des mutations particulières sont-elles alléliques ? Quelle est la fonction cellulaire d'un gène étudié ? ». En bref, l'accent est mis sur le raisonnement suivi actuellement en génétique.

LES PROBLÈMES à la fin de chaque chapitre ont été conçus pour que les étudiants appliquent et exercent leurs capacités d'analyse. Les problèmes sont présentés du plus simple au plus compliqué. Tous les problèmes ont été testés dans des classes. Un nouveau type de problème appelé *Problèmes de reconnaissance des caractéristiques observables* présente un intérêt particulier. Ces problèmes sont des représentations symboliques des résultats de systèmes de croisements simples, présentés sous forme de schémas, avec un nombre restreint de mots. Ils ont été créés pour aider les étudiants à identifier dans les données, les modes de transmission héréditaires, ce qui représente un grand avantage en analyse génétique. La plupart des chapitres comportent également un problème

Problèmes de reconnaissance de caractéristiques observables

Dans les problèmes 27 à 52, les schémas montrent les phénotypes et les résultats d'analyses de croisements. Déduisez dans la mesure du possible les génotypes des individus présentés dans chaque schéma. Tous les organismes sont diploïdes.



10. Lors d'une expérience d'interruption de conjugaison chez *E. coli*, il a été établi que le gène *pro* pénètre après le gène *thr*. Une *Hfr pro⁺ thr⁺* est croisée avec une souche *F⁻ pro⁻ thr⁻*. Les *conjugants* sont évalés sur un milieu contenant de la thiamine mais pas de proline. 360 colonies sont examinées. Elles sont isolées et cultivées sur un milieu complet. On teste ensuite chez ces cultures la capacité à pousser sur un milieu ne contenant ni proline ni thiamine (milieu minimum) et on s'aperçoit que 320 cultures seulement peuvent pousser.
- a. Déduisez les génotypes des deux types de cultures.
- b. Représentez les événements de crossing-over nécessaires pour produire ces génotypes.
- c. Calculez la distance séparant les gènes *pro* et *thr* en unités de recombinaison génétique.

Décomposons le problème

- a. Quel type d'organisme est *E. coli* ?
- b. À quoi ressemble une culture d'*E. coli* (Dense ou en une ?)
- c. Dans son habitat naturel, sur quels types de substrats pousse habituellement *E. coli* ?
- d. Quels sont les besoins élémentaires des cellules *E. coli* pour se diviser ?
- e. Définissez les termes *prototrophe* et *auxotrophe*.

contenant un exercice intitulé *Décomposons le problème*. De tels exercices révèlent les niveaux de connaissance nécessaires pour résoudre un problème de façon constructive. L'exercice décomposé évalue ces connaissances et révèle les incompréhensions fondamentales qui empêchent parfois les étudiants de résoudre les problèmes avec succès.

Chaque chapitre comporte également des **PROBLÈMES RÉSOLUS** qui montrent aux étudiants la façon dont les généticiens appliquent les principes de la génétique aux données expérimentales. Les problèmes résolus préparent les étudiants à résoudre seuls, les problèmes qui suivent.

Aide à l'étude

POUR L'ESSENTIEL, au début de chaque chapitre, donne un aperçu des principes fondamentaux traités dans le chapitre concerné, sous la forme d'un énoncé simple sans terminologie génétique. Cette rubrique constitue un résumé pédagogique utile au lecteur.

Des **MESSAGES** encadrés apparaissent au détour de chaque chapitre et constituent autant de bornes sur lesquelles le lecteur peut s'arrêter et considérer les connaissances présentées juste avant.

Chaque **RÉSUMÉ** de chapitre récapitule brièvement le contenu du chapitre concerné et permet de revoir les concepts abordés. Les résumés sont particulièrement utiles pour les révisions, en particulier pour la préparation des examens.

À la fin de chaque chapitre, il est demandé à l'étudiant d'établir un **DIAGRAMME CONCEPTUEL**. Les diagrammes conceptuels sont apparus en éducation avec le mouvement constructiviste, qui affirme que l'apprentissage des étudiants est plus efficace lorsque les dernières acquisitions sont confrontées aux connaissances antérieures. Les diagrammes conceptuels sont d'une grande efficacité pour cette confrontation et pour vérifier si les concepts ont été intégrés.

PROBLÈME RÉSOLU

2. Les feuilles d'ananas peuvent être classées en trois types : épineux (E), à bords épineux (ST) et non épineux (P). Lors de croisements entre des souches pures, suivis par des croisements entre membres de la F_2 , les résultats suivants ont été obtenus :

Croisement	Parents	Phénotypes	
		F_1	F_2
1	ST \times S	ST	96 ST : 34 S
2	P \times ST	P	120 P : 39 ST
3	P \times S	P	96 P : 25 ST : 4 S

SOLUTION

a. Premièrement, regardons les rapports de la F_2 . Nous avons des rapports 3 : 1 évidents dans les croisements 1 et 2, ce qui indique des ségrégations d'un gène unique. Le croisement 3 au contraire, présente un rapport 12 : 3 : 1.

a. Attribuez des symboles aux gènes. Expliquez ces résultats en fonction des génotypes et de leurs rapports.

b. En utilisant le modèle de la partie a, indiquez les rapports phénotypiques que vous attendriez si vous croisez (1) la descendance F_1 issue de non épineux \times épineux avec la souche parentale épineux et (2) la descendance F_1 de non épineux \times épineux avec la descendance F_1 épineux \times bords épineux.

MESSAGE

L'architecture d'une protéine est la clé de la fonction des gènes. La séquence spécifique d'acides aminés détermine la forme générale, les propriétés de liaison et la réactivité de la protéine.

RÉSUMÉ

L'élément central de la transmission des gènes au travers des générations de cellules ou d'organismes, est la réplication de l'ADN. La réplication de l'ADN précède à la fois la division cellulaire sexuelle et asexuelle. Au cours de la réplication de l'ADN, les deux moitiés de la double hélice se séparent et les simples brins servent de matrice pour une nouvelle polymérisation des nucléotides. Chaque molécule fille d'ADN est donc à moitié ancienne et à moitié nouvelle ; ce mécanisme de réplication est dit semi-conservatif. L'ADN polymérase catalyse la nouvelle polymérisation des nucléotides.

POUR L'ESSENTIEL

La prolifération cellulaire normale est modulée par la régulation du cycle cellulaire.

L'apoptose est un mécanisme normal d'auto-destruction qui élimine les cellules endommagées ou potentiellement dangereuses pour l'organisme.

Les systèmes de signalisation permettent de coordonner la prolifération et l'apoptose dans une population de cellules.

Les cellules cancéreuses prolifèrent sans contrôle et n'obéissent plus aux mécanismes vitaux de destruction, car elles ont accumulé une succession de mutations particulières dans la même cellule somatique.

Un grand nombre des classes de gènes dont la mutation provoque un cancer sont des constituants cellulaires importants qui contribuent directement ou indirectement au contrôle de la croissance des cellules et à leur différenciation.

DIAGRAMME CONCEPTUEL

Dessinez un diagramme conceptuel mettant en relation le maximum de termes parmi ceux qui sont proposés. Remarquez que ces termes ne sont pas écrits dans un ordre particulier.

Double hélice d'ADN / génome / nucléotides / homologues / pléiotropie / x / haploïde / diploïde / chromosome

Remerciements

NOUS REMERCIONS les personnes suivantes des éditions W. H. Freeman and Company pour l'aide considérable qu'elles nous ont apportée durant toute la préparation de cet ouvrage : Randi Rossignol, Sara Tenney, Nicole Folchetti, Jane O'Neill, Cambraia Magalhaes, Bill Page, Larry Marcus, ainsi que Julia DeRosa et Paul Rohloff. Nous remercions également William O'Neal et Patricia Zimmerman et le studio artistique Hudson River.

Enfin, nous tenons également à exprimer notre gratitude à nos collègues qui ont relu le manuscrit ; leurs remarques et leurs conseils nous ont été précieux :

Gwen Acton, *Université de Harvard*
Karen Anderson, *Université de Hofstra*

George Bajszar, *Université du Colorado à Colorado Springs*

Michael Bentley, *Université de Calgary*

F. Paul Doerder, *Université de l'État de Cleveland*

John Ellison, *Université du Texas A & M*

Ranjan Ganguly, *Université du Tennessee à Knoxville*

Patrick Gulick, *Université de Concord*

Nancy Hollingsworth, *Université d'État de New York à Stony Brook*

Lynne Hunter, *Université de Pittsburgh*

Richard Imberski, *Université du Maryland à College Park*

Greg Jongeward, *Université du Pacifique*

Robert Kitchin, *Université du Wyoming à Laramie*

John Locke, *Université de l'Alberta*

Joyce Maxwell, *Université d'État de Californie à Northridge*

Kim McKim, *Université de Rutgers*

Grant Mitman, *Institut universitaire technologique de l'Université du Montana*

Lenore Neigeborn, *Université de Rutgers*

Stephen Poole, *Université de Californie à Santa Barbara*

Carl Price, *Université de Rutgers*

Daniel Promislow, *Université de Géorgie*

Mary Anne Pultz, *Université de Washington Ouest*

Wm. Todd Rainey, *College Bluffton*

Dick Richardson, *Université du Texas à Austin*

John Ringo, *Université du Maine*

Mark Sanders, *Université de Californie à Davis*

Brian Shmacfsky, *College Kingwood*

Navin Sinha, *Université de Rutgers*

William Sofer, *Université de Rutgers*

Jacob Varkey, *Université d'État de Humboldt*

Glenn Yasuda, *Université de Seattle*

ANALYSE GÉNÉTIQUE MODERNE

1

LA GÉNÉTIQUE ET L'ORGANISME



POUR L'ESSENTIEL

La génétique a unifié les sciences biologiques en révélant des systèmes héréditaires homogènes.

La génétique intéresse directement les activités humaines.

La génétique peut être définie comme l'étude des gènes.

La variation génétique contribue aux variations dans la nature.

La variation génétique est le point de départ d'une grande partie de l'analyse génétique.

L'approche génétique peut être utilisée pour disséquer n'importe quel processus biologique.

Variation génétique de la couleur des grains de maïs. Chaque grain représente un individu de constitution génétique distincte. La photographie symbolise l'histoire de l'intérêt du genre humain pour l'hérédité. L'homme améliorait déjà le maïs des milliers d'années avant l'avènement de la discipline moderne que représente la génétique. C'est maintenant l'un des principaux organismes étudiés en génétique classique et en génétique moléculaire.

(William Sheridan, Université du Dakota du Nord ; photo de Travis Amos.)

Pourquoi étudier la génétique ? Il y a deux raisons fondamentales à cela. La première, c'est que la génétique est devenue une plaque tournante de la biologie et que comprendre la génétique est capital pour tous ceux qui désirent étudier à fond la vie des plantes, des animaux ou des micro-organismes. La seconde, c'est que la génétique, plus que toute autre discipline scientifique, occupe aujourd'hui une position centrale dans divers secteurs des activités humaines et, à ce titre, concerne le genre humain de bien des façons. En effet, des questions liées à la génétique se posent quotidiennement et il n'est plus possible de prétendre en ignorer les découvertes. Ce chapitre présente un panorama de la génétique en tant que science et montre comment elle en est arrivée à occuper cette position essentielle. Nous y exposons en outre la manière d'aborder les chapitres suivants.

Commençons par définir la génétique. Certains la présentent comme l'étude de l'hérédité, mais les phénomènes héréditaires ont intéressé l'homme longtemps avant la naissance de notre civilisation. Bien avant que la biologie ou la génétique ne constituent les disciplines scientifiques que nous connaissons aujourd'hui, les peuples anciens amélioraient déjà les plantes cultivées et les animaux domestiques en sélectionnant des individus qu'ils réservaient à la reproduction. Ils ont certainement été intrigués par la transmission des caractères chez l'homme et se sont demandés « pourquoi les enfants ressemblent à leurs parents » ou « comment certaines maladies se transmettent dans les familles ». Cependant, la génétique en tant qu'ensemble de principes et de méthodes analytiques débuta seulement dans les années 1860 lorsqu'un moine augustin nommé Gregor Mendel (figure 1-1) réalisa une série d'expériences qui mirent en évidence des éléments biologiques appelés **gènes**. Le mot *génétique* vient de « gènes » et les gènes constituent l'objet principal de cette discipline. Que les généticiens mènent leurs études au niveau de la molécule, de la cellule, de l'organisme, de la population ou encore de l'évolution, les gènes sont toujours au centre de leurs préoccupations. En termes simples, la génétique est l'étude des gènes.

Que sont les gènes ? Les gènes sont des segments d'une macromolécule filiforme organisée en double hélice, l'**acide désoxyribonucléique**, dont le nom est abrégé en **ADN**. L'ADN, le matériel héréditaire qui est



Figure 1-1 Gregor Mendel. (Musée Morave, Brno.)

transmis d'une génération à la suivante, détermine les propriétés intrinsèques d'une espèce. L'information codée dans l'ADN a la forme d'une succession de sous-unités, appelées **nucléotides**. Chacune des cellules d'un organisme contient un ou deux exemplaires d'un jeu d'ADN appelé un **génome**. Le génome comporte une ou plusieurs molécules d'ADN extrêmement longues, organisées en structures que l'on appelle des **chromosomes**. Les gènes sont simplement les unités fonctionnelles de l'ADN chromosomique. Chaque gène code non seulement la structure d'un produit cellulaire, mais porte également des éléments de contrôle qui déterminent quand, où et combien de produit doit être synthétisé (figure 1-2). La plupart des gènes codent des produits protéiques (protéines). Les protéines sont les déterminants les plus

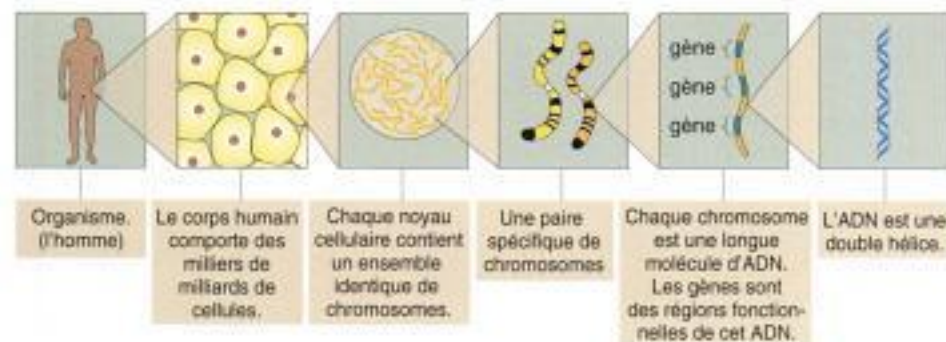


Figure 1-2 Agrandissements successifs de l'organisme au matériel génétique.

importants des propriétés des cellules et des organismes : lorsque vous regardez un organisme, ce que vous voyez est soit une protéine, soit quelque chose qui a été fabriqué par une protéine. Les gènes codent les protéines de façon indirecte et en plusieurs étapes. La première étape consiste à copier (**transcrire**) l'information codée dans l'ADN du gène, sous la forme d'une molécule simple-brin apparentée à l'ADN, appelée **ARN messager**. Ensuite l'information contenue dans l'ARN messager est **traduite** (décodée) en une chaîne d'acides aminés appelée un **polypeptide**. Les polypeptides, seuls ou en association avec d'autres polypeptides et constituants cellulaires, forment les **protéines** fonctionnelles de la cellule. Puisque les propriétés des cellules déterminent les propriétés d'un organisme, les gènes ont un lien direct avec la structure et la fonction d'un organisme. Chez les organismes les plus complexes, le nombre de chromosomes est généralement de l'ordre de la dizaine tandis que le nombre de gènes est de l'ordre de plusieurs dizaines de milliers.

Armés de ces définitions de la génétique et des gènes, voyons maintenant pourquoi la compréhension de ces notions a pris une telle importance. Nous envisagerons d'abord l'impact de la génétique sur notre propre espèce.

LA GÉNÉTIQUE ET LES ACTIVITÉS HUMAINES

La génétique semble occuper une place particulière dans les activités humaines. Elle n'est pas seulement importante en tant que discipline scientifique, mais elle l'est aussi parce qu'elle nous apprend énormément sur la nature du genre humain. Elle occupe à ce titre une position unique parmi les sciences biologiques.

La société moderne a besoin de la génétique

Jetez un coup d'œil aux vêtements que vous portez. Le coton de votre chemise et de votre pantalon provient de cotonniers qui sont différents de leurs ancêtres sauvages. Ils résultent en effet de programmes intensifs d'amélioration, basés sur l'application systématique des principes de la génétique. Il en est de même des moutons qui ont produit la laine de votre chandail et de votre veston. Songez aussi à vos derniers repas. Vous pouvez être certain que le riz, le blé, le poulet, le porc et les principaux autres organismes qui servent à alimenter les habitants de cette planète ont été spécialement croisés et sélectionnés grâce à l'application de procédés génétiques courants, par des cultivateurs et des éleveurs (figure 1.3). La « Révolution verte », qui a si spectaculairement accru la productivité globale des plantes cultivées, est un succès de la génétique qui a permis de mettre au point des lignées hautement productives pour certaines espèces essentielles (figure 1-4).

Pour obtenir une production agricole maximum, un fermier peut semer des graines de la même constitution génétique sur une vaste superficie ; c'est la pratique de la



Figure 1-3 Un généticien réalisant des croisements contrôlés entre des plants de maïs. Les sacs en papier servent à éviter des pollinisations croisées accidentelles et à récolter le pollen qui servira aux croisements spécifiques désirés. (Chuck O'Rear / Woodfin Camp.)

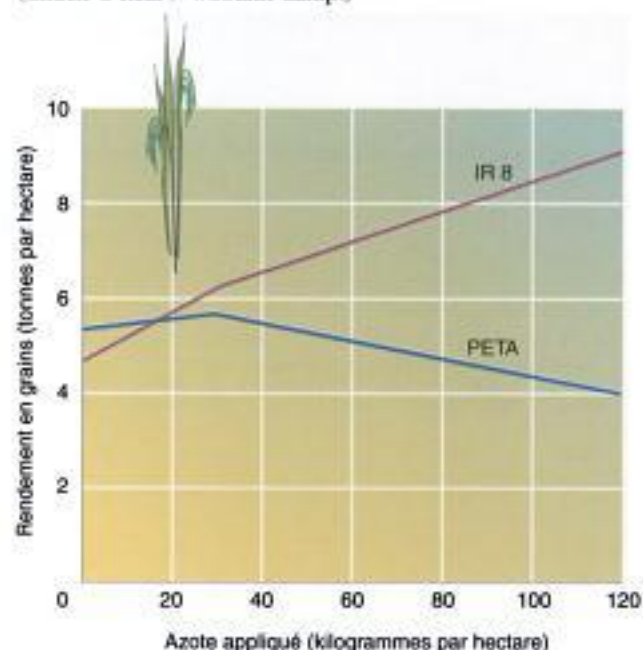


Figure 1-4 La souche améliorée de riz nain appelée *IR 8* doit en partie son succès à sa remarquable réponse à l'application d'engrais. PETA, une souche plus ancienne de taille normale, montre une réponse plus typique. (D'après Peter R. Jennings, « The amplification of agricultural production. » *Scientific American*, Septembre 1976, p 180-194.)

monoculture. Puisque des organismes pathogènes (les organismes qui provoquent des maladies) ou des prédateurs s'attaquent aux cultures, des gènes de résistance déterminés sont introduits par croisement dans les lignées cultivées pour les protéger. Dans certains cas, les gènes de résistance sont efficaces pendant des décennies, mais dans d'autres, les pathogènes triomphent des gènes de résistance en quelques années. Des modifications génétiques aléatoires se produisent en permanence dans les populations d'organismes pathogènes et peuvent parfois leur conférer des capacités pathogènes nouvelles. Ces modifications mettent en péril la monoculture tout entière. C'est pourquoi les généticiens des plantes doivent toujours avoir une longueur d'avance sur les pathogènes afin d'éviter les épidémies ou la destruction des végétaux et leurs conséquences catastrophiques sur la production des produits alimentaires (tableau 1-1 et figure 1-5).

Les champignons et les bactéries ont également été améliorés pour les besoins de l'homme. Les levures en sont un exemple manifeste. Elles sont à la base d'industries plusieurs fois milliardaires qui fabriquent des produits de boulangerie, des boissons alcoolisées ou encore de l'alcool qui sert de carburant. Les champignons produisent la pénicilline, un antibiotique, ou la cyclosporine, un immunosuppresseur qui évite le rejet des greffes d'organes, ainsi que toute une gamme de produits industriels importants comme l'acide citrique et l'amylase. Les bactéries fournissent d'autres antibiotiques importants pour les sciences médicales comme la streptomycine. La plupart des industries qui utilisent des champignons et des bactéries ont bénéficié de l'application des principes de la génétique classique, mais nous sommes à présent dans une ère nouvelle où les scientifiques peuvent manipuler direc-

tement l'ADN pour élaborer des souches microbiennes entièrement nouvelles qui sont conçues exclusivement en tube à essai pour satisfaire les besoins de l'homme. Cette technologie s'appelle le **génie génétique moléculaire**. C'est de cette manière par exemple, qu'on dispose de souches bactériennes qui produisent des protéines de mammifères comme l'insuline, utilisée pour le traitement du diabète, ou l'hormone de croissance qui sert à soigner le nanisme hypophysaire. L'insuline produite par ces souches est la véritable hormone humaine, contrairement à celles que l'on utilisait antérieurement, qui étaient d'origine bovine ou porcine. Certaines des substances produites par ces souches microbiennes sont extrêmement difficiles à produire à partir d'autres sources : par exemple, la quantité d'hormone de croissance requise pour le traitement d'un enfant pendant un an exigeait auparavant l'extraction des hypophyses de 75 cadavres humains. En fait, certaines souches spéciales de champignons et de bactéries sont manipulées génétiquement pour accroître les quantités des produits utiles qu'on en extrait. Le génie génétique moléculaire fut appliqué en premier lieu aux micro-organismes, mais à l'heure actuelle les mêmes techniques sont appliquées aux plantes et aux animaux et fournissent des lignées manipulées qui n'auraient jamais pu être obtenues par croisements sélectifs.

Le génie génétique moléculaire consiste à modifier un génome en exposant des cellules à des fragments d'ADN « étranger » portant des gènes qui codent les protéines désirées, provenant souvent d'une autre espèce. Cet ADN est absorbé par les cellules puis il s'insère dans un ou plusieurs chromosomes receveurs ; il est alors transmis comme n'importe quel autre gène. Les cellules modifiées de la sorte sont dites **transgéniques**. À partir d'une cellule

TABLEAU 1.1
Résistance des souches de riz aux maladies et aux insectes.

Souches	Année de développement	MALADIES				INSECTES			
		<i>Pyricularia oryzae</i> (ascomycète)	Rouille bactérienne	Virus des striures foliaires	Virus du rabougrissement herbeux	Les virus du tungro du riz	Sauterelle verte	Criquet	Insecte mineur
IR 8	1966	MR	S	S	S	S	R	S	MS
IR 5	1967	S	S	MS	S	S	R	S	S
IR 20	1969	MR	R	MR	S	R	R	S	MS
IR 22	1969	S	R	MS	S	S	S	S	S
IR 24	1971	S	S	MR	S	MR	R	S	S
IR 26	1973	MR	R	MR	MR	R	R	R	MR

NOTE : Les entrées décrivent la sensibilité ou la résistance de chaque souche à chaque maladie ou insecte de la manière suivante : S = sensible ; MS = modérément sensible ; MR = modérément résistante ; R = résistante. (D'après *Research Highlights*, International Rice Research Institution, 1973, p. 11.)



Figure 1-5 Des lignées pures de différentes espèces de riz sélectionnées pour leur résistance à divers parasites des végétaux. (Patrick Aventurier / Gamma Liaison.)

transgénique, on peut produire un organisme transgénique dont toutes les cellules contiendront cet ADN étranger supplémentaire. Grâce à cette technique, on peut produire une vaste gamme d'organismes économiquement rentables. L'application commerciale du génie génétique est généralement appelée la **biotechnologie**. On estime que l'industrie biotechnologique basée sur la génétique deviendra dans les prochaines décennies l'une des industries les plus rentables financièrement. Quoi qu'il en soit, la biotechnologie est un des éléments de la source d'abondance que représente la génétique, en contribuant à la qualité de vie de l'importante population mondiale actuelle.

Un autre secteur des activités humaines qui fait appel à la génétique est celui de la médecine légale. Tous les êtres humains possèdent plus ou moins le même assortiment d'ADN. Mais des variations mineures de cet ADN sont responsables des variations entre les hommes. Si l'on dispose d'un échantillon d'ADN de taille conséquente provenant d'un individu, on peut lui appliquer des techniques particulières qui permettent d'examiner un grand nombre de segments d'ADN et d'en tirer une empreinte d'ADN (*fingerprint*). Sauf dans le cas de vrais jumeaux, de telles empreintes sont spécifiques de chaque individu. Les empreintes d'ADN sont largement utilisées dans les enquêtes policières. Même des empreintes d'ADN végétal peuvent être utiles. Ainsi, on a rapporté un cas où l'ADN de deux minuscules graines trouvées à l'arrière d'un camion suspect a permis d'établir très clairement que celui-ci avait été garé sous un arbre à proximité duquel le corps de la victime d'un meurtre avait été retrouvé. On ne dispose souvent que de très petits échantillons d'ADN provenant de sang, de sperme ou même de follicules pileux. Les segments de l'empreinte doivent donc être étudiés un à un. Comme certains fragments peuvent être communs à un grand nombre d'individus, ce genre de preuves doit être interprété avec davantage de prudence. Ces empreintes partielles peuvent permettre d'éliminer ou de diminuer le nombre de suspects possibles dans une affaire criminelle, mais ne peuvent servir à identifier un individu précis.

La génétique intervient de manière essentielle en médecine

Beaucoup de maladies humaines ont une origine génétique. On estime, par exemple, qu'au moins 30 pour cent des admissions en hôpital pédiatrique impliquent une composante génétique. Mais les recherches actuelles montrent de plus en plus l'importance des prédispositions d'origine génétique dans des maladies graves aussi bien que dans des affections plus bénignes, de telle sorte que ces chiffres sont presque certainement sous-estimés.

Les maladies génétiques présentent de multiples facettes. Un grand nombre des maladies génétiques héréditaires (telles que la mucoviscidose, la phénylcétonurie et la dystrophie musculaire) sont provoquées par des formes anormales, des **mutations**, de gènes spécifiques transmis par les gamètes (spermatozoïde et ovule). Les positions sur les cartes des chromosomes humains, de ces gènes dont les formes anormales sont responsables des maladies héréditaires les plus connues, sont indiquées dans la figure 1-6 à la page suivante. Un grand nombre des gènes présentés ont été isolés et analysés au niveau moléculaire. Les gènes responsables des formes héréditaires de la maladie d'Alzheimer et du cancer du sein ont été isolés récemment. Les formes anormales du gène *S182* localisé sur le chromosome 14 sont responsables de 80 pour cent des maladies d'Alzheimer héréditaires précoces soit 10 pour cent de tous les cas de cette maladie. Le gène *S182* code normalement une protéine membranaire. Les formes héréditaires du cancer du sein représentent 5 pour cent de tous les cancers du sein de la population nord-américaine. Les formes anormales des gènes *BRCA1* et *BRCA2* sont la cause de deux tiers des formes familiales de cancer du sein. *BRCA1* est également associé à une prédisposition aux cancers ovariens. Ce gène code probablement une protéine régulant l'activité d'autres gènes. Dans les deux cas, comprendre de quelle façon ces gènes provoquent ces formes héréditaires rares de la maladie conduira certainement à une thérapie efficace et à une meilleure compréhension des formes « sporadiques » de ces maladies.

Le déterminisme génétique de certaines affections héréditaires est complexe car elles sont causées par l'interaction de plusieurs gènes qui interagissent également tous avec l'environnement. Les affections cardiaques et les diabètes en sont des exemples. Mais la technologie de la génétique moléculaire a ouvert des voies nouvelles pour découvrir et isoler ces gènes à interactions multiples.

L'isolement de gènes d'animaux de laboratoire peut également conduire à des progrès dans le domaine de la santé. On a par exemple isolé des gènes uniques qui ont un effet important sur l'obésité des souris. Un de ces gènes, *ob*, code une protéine de signalisation qui appartient au système indiquant à la souris qu'elle a absorbé suffisamment d'aliments. Un autre gène, *CPE*, induit lentement l'obésité chez la souris, qui présente alors une prédisposition au diabète.

Le type le plus courant de cancer « sporadique » est provoqué par l'apparition soudaine d'une forme mutante

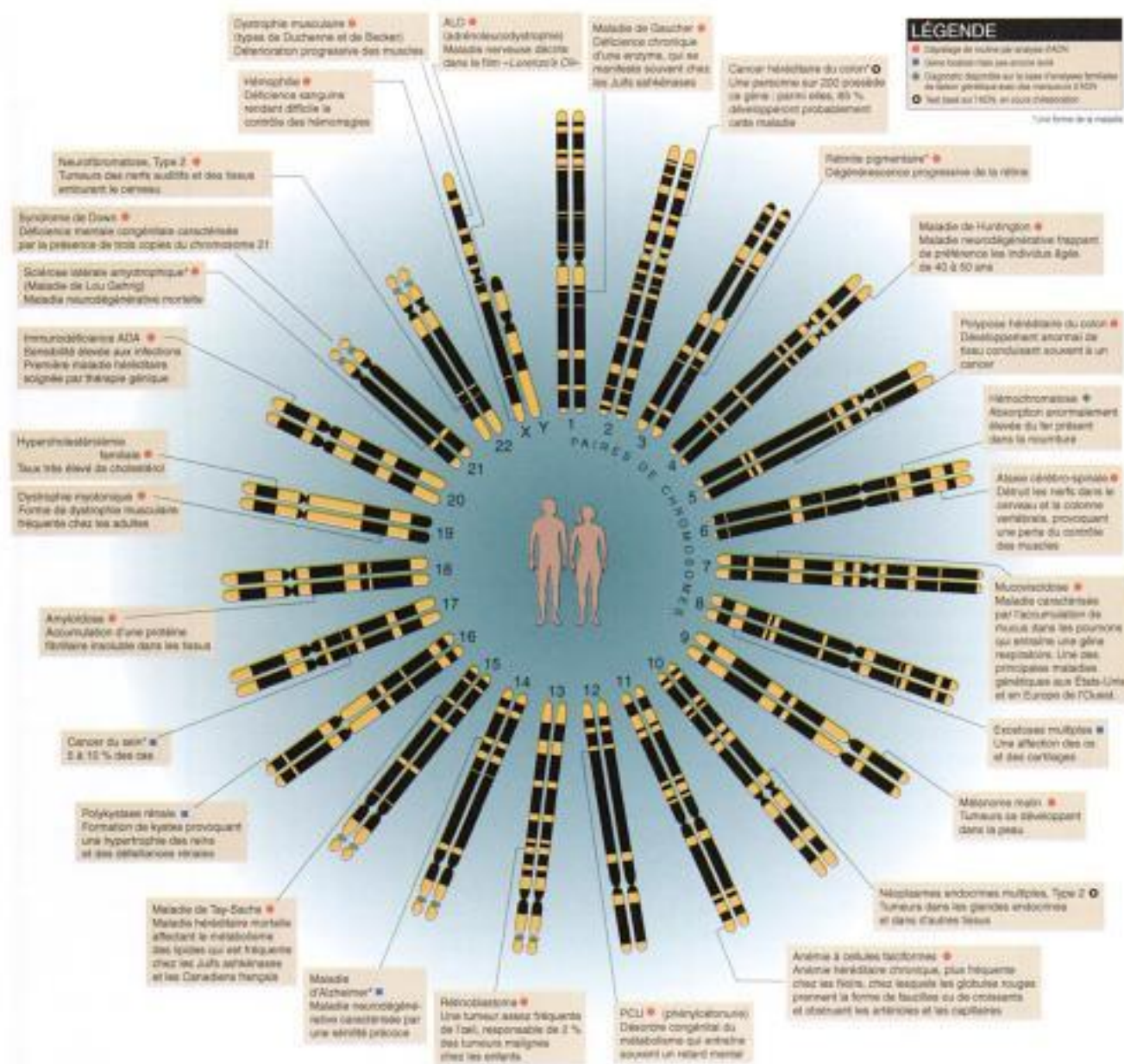


Figure 1-6 Les 23 chromosomes humains avec les positions des gènes dont les formes anormales sont responsables des maladies héréditaires les plus connues. (Time.)

d'un gène dans une cellule *somatique*, c'est-à-dire une cellule du corps qui ne produit ni ovule ni spermatozoïde. Puisque ce type de cancers se produit dans des tissus somatiques, par définition, ils ne sont pas transmis à la descendance. Cependant, comme nous l'avons vu précédemment, la prédisposition à certains cancers est héréditaire. On est assez peu conscient du fait que le cancer, qui nous touche tous dans notre vie d'une façon ou d'une autre, est une maladie génétique.

Enfin, un autre type de maladie génétique est dû à des **aberrations chromosomiques**, anomalies de la

structure ou du nombre de chromosomes. Le syndrome de Down en est un exemple commun ; il est provoqué par la présence d'une copie supplémentaire d'un chromosome spécifique. La plupart des cas de syndrome de Down apparaissent *de novo*, à chaque génération, à la suite d'accidents cellulaires qui se produisent dans les gonades de parents qui possèdent par ailleurs un assortiment chromosomique parfaitement normal. De tels accidents conduisent à un spermatozoïde ou un ovule anormal, ce qui engendre alors un enfant atteint du syndrome de Down.

La génétique a montré que les gènes sont à l'origine de nombreuses maladies, mais très souvent, c'est elle aussi qui permet de soulager des souffrances. On utilise déjà des sondes moléculaires pour détecter les gènes défectueux chez les futurs parents. Et il est même possible d'isoler et de caractériser ces gènes défectueux grâce à des techniques de génétique moléculaire. Par exemple, l'ADN du gène qui cause la mucoviscidose a été isolé et caractérisé. En découvrant la séquence nucléotidique d'un gène responsable d'une maladie, il est possible de déduire la structure et la fonction de la protéine codée par le gène et donc de comprendre le défaut physiologique provoqué par la mutation. Une fois la maladie comprise, on est en mesure de concevoir de meilleurs traitements thérapeutiques. En fin de compte, on peut espérer que la **thérapie génique** directe guérira de nombreuses maladies génétiques. La thérapie génique consiste à insérer la copie normale d'un gène dans des cellules qui contiennent son homologue défectueux. Les premiers cas de thérapie génique humaine décrits, concernent deux enfants souffrant d'un déficit immunitaire sévère combiné (une affection héréditaire des globules blancs). Des cellules furent prélevées sur les patients, un gène fonctionnant correctement y fut inséré, et les cellules furent réinjectées aux enfants pour être incorporées dans la population des cellules sanguines de leur organisme. Depuis ce premier cas, bien d'autres essais de thérapie génique ont été rapportés et il y en aura sans conteste de plus en plus dans les années à venir. Quoi qu'il en soit, la thérapie génique en est encore à ses balbutiements. L'action d'un seul gène dépend de nombreux facteurs, qui doivent tous fonctionner de concert pour supprimer les symptômes des maladies traitées.

Les généticiens étudient également le virus de l'immunodéficience humaine (VIH) qui provoque le syndrome de l'immunodéficience acquise (SIDA). Au cours de leur cycle normal de reproduction, les virus du type VIH insèrent des copies de leur matériel génétique dans les chromosomes des individus qu'ils infectent. C'est pourquoi, dans un sens, le SIDA est également une maladie génétique. La détermination de la manière dont ces gènes viraux s'intègrent et fonctionnent constitue par conséquent une étape importante pour vaincre la maladie.

Un projet de recherche considérable est en cours. Il s'agit du séquençage des trois milliards de nucléotides du génome humain. La médecine en sera le principal bénéficiaire. Ce projet « Génome Humain » est une collaboration internationale impliquant de nombreux laboratoires, chacun travaillant au séquençage de régions chromosomiques spécifiques. Des cartes physiques de haute résolution ont déjà été obtenues pour l'ADN complet de plusieurs chromosomes (dont le chromosome Y et le chromosome 21). Ces cartes sont un préalable nécessaire au séquençage complet des chromosomes. Les bénéfices potentiels d'un tel projet sont énormes pour la médecine. Mais plus passionnante encore est la perspective de connaître le plan d'assemblage de ce que l'on peut considérer comme la structure la plus complexe de l'univers, l'*Homo sapiens* (figure 1-7).



Figure 1-7 Deux biologistes moléculaires analysant des séquences d'ADN. Les bandes sombres représentent des photographies de fragments radioactifs d'ADN. (David Parker / Science Source / Photo Researchers.)

La génétique influence notre vision du monde

Chacun de nous acquiert graduellement, depuis le moment où il devient conscient, une vision personnelle de l'univers et de la place qu'il y occupe. Cette perception du monde, représentative de notre identité en tant qu'individu, nous inspire nos attitudes et notre manière d'agir. Elle détermine le type d'homme ou de femme que nous sommes et, en fin de compte, la société dans laquelle nous vivons. Toute connaissance nouvellement acquise doit s'intégrer dans cette vision de l'univers à moins de modifier celle-ci pour l'adapter à ce nouvel élément. La génétique a donné naissance à un certain nombre de concepts nouveaux qui ont changé radicalement la manière dont l'homme se perçoit et se situe par rapport au reste de l'univers.

Le meilleur exemple de changement de notre vision du monde vient sans doute des travaux de génétique, des études des chromosomes, de l'ADN et des protéines, qui montrent que nous sommes non seulement apparentés aux singes et aux autres mammifères mais aussi, ce qui est plus inattendu, à tous les autres êtres vivants de la planète, y compris les plantes, les champignons et les bactéries. Tous les êtres vivants possèdent le même système de stockage et d'expression de l'information génétique et présentent une homologie (ressemblance basée sur la parenté dans l'évolution) de nombreuses structures y compris au niveau des gènes. Le fait qu'il existe une parenté dans le monde vivant est une notion importante parce qu'elle

nous intègre parmi les autres êtres vivants. Cette notion modifie radicalement notre vision du monde puisque le genre humain ne peut plus être considéré comme le pî-nacle ou le centre de la création mais comme une forme vivante égale aux autres. Il faut le reconnaître, ceci relève de la philosophie et de la religion mais c'est précisément le fond du problème : la génétique nous force à reconsidérer la manière dont nous nous percevons.

Certaines des questions sociales les plus préoccupantes impliquent une composante génétique indirecte. Les problèmes d'injustice et de préjugés sociaux les plus graves posent des questions sur les différences de comportement entre les populations et entre les sexes. Le QI est-il lié à l'appartenance à une population ? Les hommes et les femmes ont-ils des prédispositions génétiques différentes à un comportement agressif ? La génétique est un moyen d'analyser et de réfléchir à ces questions complexes qui attendent encore une réponse.

Une des préoccupations essentielles des biologistes est la vitesse alarmante à laquelle nous détruisons les habitats naturels, en particulier sous les tropiques, qui abritent de vastes réservoirs de vie animale et végétale. Là encore, le problème présente une composante génétique essentielle puisque la question est la sauvegarde de la diversité génétique. Il est nécessaire de connaître le niveau de diversité minimum indispensable aux populations en bonne santé et la taille des populations qui permet de maintenir ce niveau. La compréhension réelle de l'impact de la destruction des habitats naturels exige de comprendre la génétique.

La santé génétique de nos populations est un autre domaine où la génétique peut avoir des retombées importantes. De nombreux généticiens s'en préoccupent car le génome humain est exposé aux agressions d'un nombre croissant d'agents de l'environnement, en particulier des radiations et des substances chimiques, susceptibles de provoquer des modifications aléatoires dans les gènes. La grande majorité de ces modifications sont probablement nuisibles. À court terme, elles n'influenceront sans doute pas significativement la fréquence des maladies héréditaires mais elles pourraient s'accumuler et, à plus longue échéance, se combiner et interagir pour produire des maladies, un effet qui a été appelé « bombe à retardement génétique ».

Bien sûr, comme tout progrès technologique, la technologie génétique apporte sa part de problèmes éthiques. La « Révolution verte », par exemple, est un succès en termes de productivité mais les concentrations élevées d'engrais et de pesticides exigées par les nouvelles variétés végétales suscitent des inquiétudes en ce qui concerne la pollution de l'eau et le coût élevé des engrais pour les agriculteurs des pays pauvres. La possibilité de créer deux brebis génétiquement identiques (de cloner une brebis — voir Les grandes découvertes de la génétique 1-1) a suscité un débat enflammé pour savoir s'il était moral d'appliquer la même technique à l'homme. En outre, la synthèse de nouvelles formes de vie transgéniques par génie génétique a soulevé de grandes inquiétudes sur la moralité et la légalité de créer des végétaux et des animaux modifiés. L'accès aux informations qui concernent la

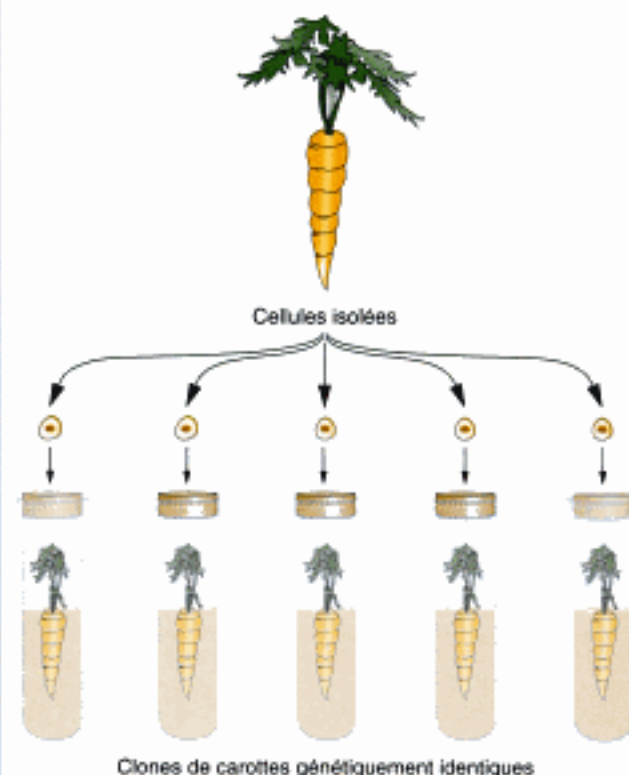
constitution génétique des individus peut également être un problème dans des domaines comme l'assurance maladie ou la médecine légale. On a déjà publié des éléments crédibles sur l'existence de gènes qui prédisposeraient à des comportements agressifs ou à l'homosexualité. Quel serait l'impact de la connaissance de ces gènes sur les droits des individus qui en sont porteurs ? Il serait fort désinvolte de dire que ce n'est pas le problème des scientifiques mais celui de la société ; les scientifiques doivent partager la responsabilité des conséquences de leurs découvertes pour la société et en envisager minutieusement toutes les retombées avant de les publier.

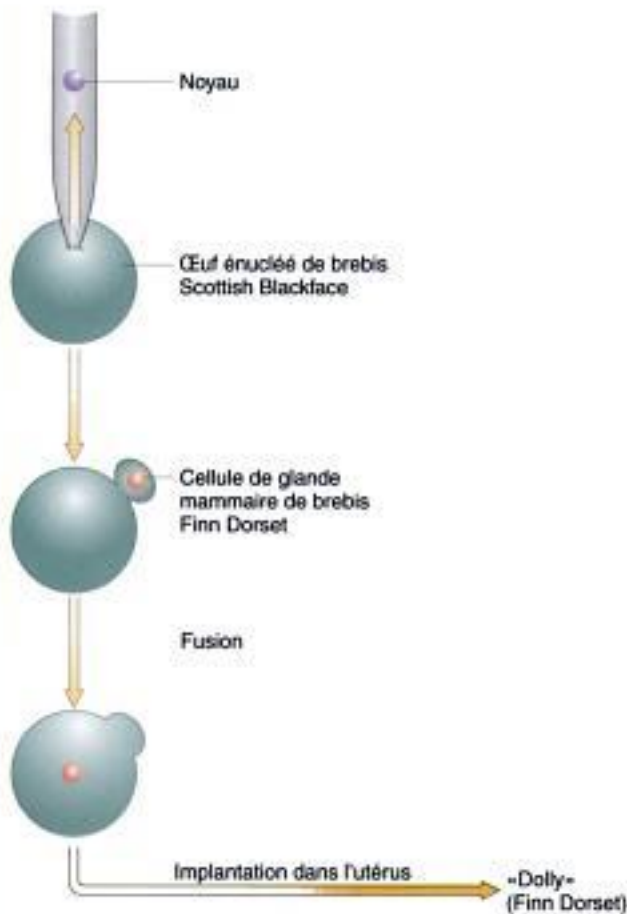
Les exemples cités ci-dessus montrent que nombre des problèmes auxquels nous sommes confrontés quotidiennement en tant qu'individus ou en tant que société sont intimement liés à la génétique. Ce n'est que par la compréhension des aspects génétiques de ces problèmes que nous pouvons espérer prendre des décisions adéquates en ce qui concerne le monde de plus en plus complexe et peuplé auquel nous appartenons.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 1-1

Fabriquer des mammifères génétiquement identiques

Les mammifères génétiquement identiques — en particulier les humains — sont un sujet d'intérêt considérable tant pour les





génétiens que pour les non-scientifiques. Les vrais jumeaux, qui apparaissent à la suite de la scission en deux d'un œuf fécondé pour créer deux individus, ont toujours été regardés comme des objets de curiosité et ont même été à la base d'essais scientifiques pour évaluer la composante génétique de la variation des phénotypes humains, tels que certains types de comportement.

Il y a plus de vingt ans, on a découvert que des cellules végétales isolées pouvaient croître en culture et produire un ensemble de végétaux génétiquement identiques, appelé un clone. Au même moment, on a réussi à cloner le crapaud africain *Xenopus laevis* en utilisant une micropipette pour enlever des noyaux de cellules intestinales de têtard et les injecter dans des œufs de crapaud énucléés. À l'évidence, les cellules de têtard avancées dans leur développement contenaient toute l'information génétique nécessaire au développement normal d'un animal adulte, un état appelé *totipotence*. Toutefois, les stades ultérieurs de développement n'étaient pas totipotents.

En février 1997, la Une des journaux du monde entier présentait la photographie d'une sympathique brebis nommée Dolly, qui fut créée en implantant le noyau d'une cellule de glande mammaire de brebis Finn Dorset dans l'œuf énucléé d'une brebis Scottish Blackface. La procédure apparemment

simple ne réussissait que si les cellules donneuses étaient au préalable cultivées en milieu nutritif pauvre afin d'induire l'arrêt de la division mitotique, une ruse qui semblait rendre le noyau donneur compatible avec l'œuf receveur, qui se trouvait à un stade tardif de la méiose. Ce résultat expérimental surprenant a montré que le noyau de mammifère adulte est toujours totipotent.

L'histoire a fait les gros titres, non en raison de cette totipotence, mais parce qu'il était évident que les développements de la technologie qui a produit Dolly pouvaient également être utilisés pour cloner des êtres humains génétiquement identiques. Ceci a soulevé un débat passionné sur l'éthique d'une telle procédure et sur ses détournements possibles. Plusieurs pays ont alors établi des lois qui réglementent ou même interdisent ce type d'expérience sur l'homme. Y a-t-il des utilisations légitimes du clonage humain ? On a suggéré par exemple que lorsqu'un bébé ou un jeune enfant mourait à la suite d'un accident, cette existence brève pourrait être prolongée par clonage. Cependant, actuellement, la question est en suspens car l'efficacité du processus de clonage de mammifère est toujours faible — de nombreux embryons défectueux ont été produits avant Dolly.



À gauche : Dolly (Finn Dorset). À droite : La mère de Dolly (Scottish Blackface). (Aimablement communiqué par l'Institut Roslin, Édimbourg, Écosse.)

LA GÉNÉTIQUE ET LA BIOLOGIE

La biologie est un domaine immense. La planète Terre abrite un nombre stupéfiant de formes vivantes. On connaît déjà l'existence de 286 000 espèces de plantes à fleurs, de 500 000 espèces de champignons et de 750 000 espèces d'insectes. Il en reste en outre un grand nombre à découvrir. Il y a cinquante ans, la biologie était divisée en disciplines distinctes, chacune étudiant la vie à un niveau différent. On distinguait la morphologie, la physiologie, la biochimie, la taxonomie, l'écologie, la génétique, etc., qui se cantonnaient dans des compartiments pratiquement isolés. Or, les découvertes de la génétique ont fourni des thèmes unificateurs à l'ensemble de la biologie, de sorte qu'aujourd'hui des liens conceptuels relient toutes ces disciplines.

Le principal lien thématique offert par la génétique est littéralement un fil, puisqu'il s'agit de la molécule d'ADN. Nous savons aujourd'hui que l'ADN est *l'information à la base de tous les processus et structures de la vie*. La structure de la molécule d'ADN rend compte de deux des caractéristiques majeures de la vie, la *reproduction* et la *genèse de la forme*. Nous verrons dans ce livre que l'ADN a une structure en double hélice, ce qui lui permet d'être répliqué sous la forme de deux copies identiques. La répllication de l'ADN est à la base de toute reproduction, tant au niveau de la cellule que de l'organisme. L'ADN est répliqué avant la division cellulaire, ce qui permet aux chromosomes de se diviser en **chromatides** qui deviennent ensuite des **chromosomes fils** qui seront transmis aux **cellules filles**. Ce processus de répllication et de formation de chromatides est fondamentalement semblable pour la division des cellules sexuelles et non sexuelles ; il est schématisé dans la figure 1-8. (Notons cependant que les deux types de division cellulaire présentent de nombreuses différences que nous passerons en revue dans des chapitres ultérieurs). C'est ce processus de répllication de l'ADN qui permet d'effectuer des copies des cellules et des organismes et de les perpétuer (figure 1-9). L'ADN peut donc être considéré comme le fil qui nous rattache à nos ancêtres dans l'évolution. En outre, l'ADN est responsable de la forme des organismes et de leurs constituants. En effet, la séquence linéaire des éléments de construction d'une molécule d'ADN contient un code qui porte les instructions nécessaires à l'élaboration d'un organisme. Nous pouvons considérer cela comme de l'information ou comme « ce qui est nécessaire pour déterminer la forme ». Les caractéristiques d'une espèce, qu'il s'agisse de sa structure ou de ses processus physiologiques, sont sous l'influence de son ADN. Dès lors, sous-jacent aux structures observées par les morphologistes, aux réactions étudiées par les physiologistes, ou aux ressemblances constatées par les évolutionnistes, nous percevons le fil unificateur de la molécule d'ADN.

L'ADN fonctionne quasiment de la même façon chez tous les organismes. En soi, ceci constitue un autre thème unificateur, mais cela signifie également que ce que nous apprenons pour un organisme a souvent des principes applicables à d'autres. C'est pour cette raison que la génétique a souvent fait un usage intensif d'**organismes modèles** dont beaucoup seront évoqués dans la suite de cet ouvrage. En réalité, le développement de la génétique hu-

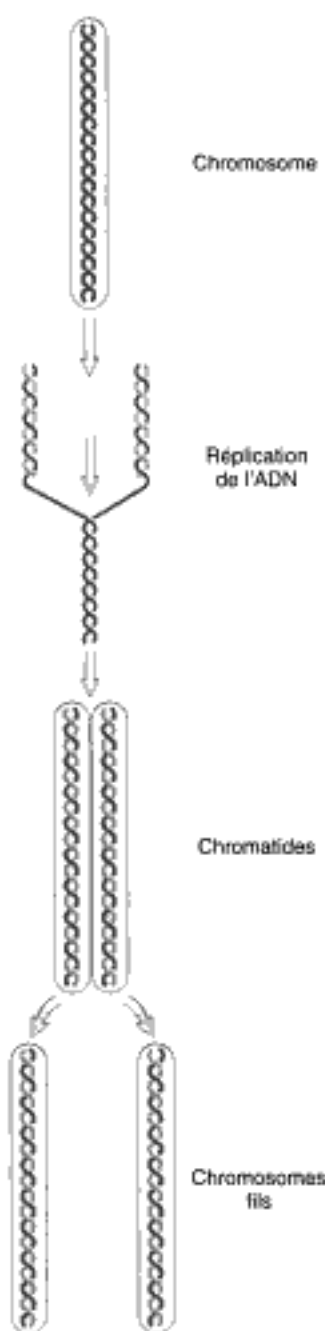


Figure 1-8 Lorsque de nouvelles cellules sont fabriquées, la répllication de l'ADN permet à un chromosome de se transformer en une paire de chromatides, qui deviendront à la fin du processus les chromosomes fils et seront transmis aux nouvelles cellules.

maine au cours des dernières décennies fut dans une large mesure permis par les progrès réalisés dans l'étude d'organismes modèles peu complexes tels que les bactéries et les champignons.

La génétique a également fourni certaines des approches analytiques les plus fines utilisées actuellement dans tout l'éventail des disciplines biologiques. Au tout premier rang se trouve la **dissection génétique**. Par cette approche expérimentale, toute structure ou tout processus

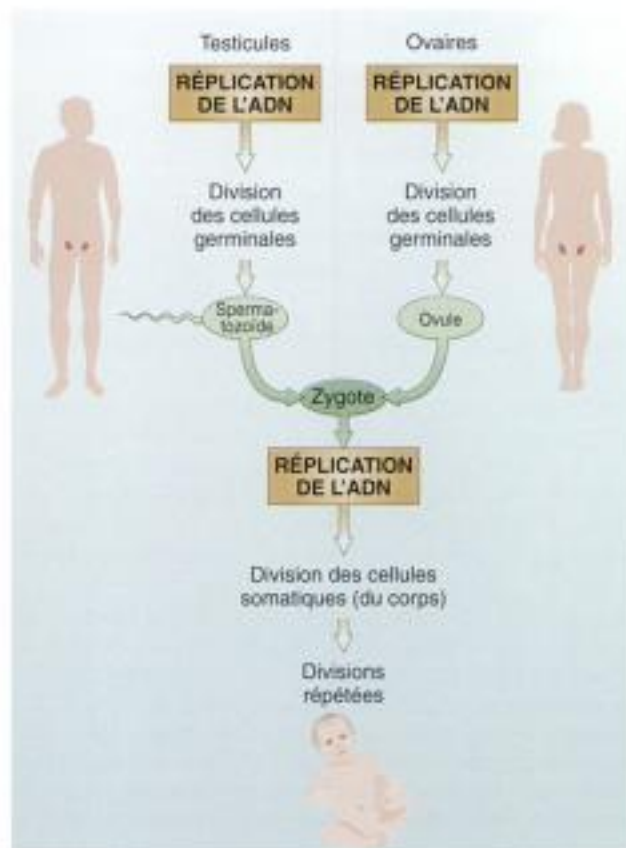


Figure 1-9 La réplication de l'ADN est la base de la transmission de la vie au travers des âges.

peut être « disséqué » en observant la façon dont les gènes *mutants* influencent. En étudiant l'anormalité, on peut en déduire la situation normale. Ainsi, dans l'étude du développement d'organismes adultes à partir d'œufs fécondés, chacun des gènes mutants qui conduit à une anomalie du développement permet d'identifier un composant du déroulement normal de celui-ci. La dissection génétique des souches mutantes paralysées de vers nématodes a permis de caractériser les gènes qui contrôlent le mouvement *normal*. Le fonctionnement complet d'un processus particulier peut être déterminé en mettant en relation tous ces composants contrôlés génétiquement.

Nous avons vu que le génie génétique moléculaire a ouvert de nouvelles perspectives en biotechnologie appliquée mais les mêmes techniques sont tout aussi utiles en recherche fondamentale. Les scientifiques ont manipulé des gènes de levure pour produire des chromosomes fonctionnels entièrement artificiels, capables de porter de grandes quantités d'ADN étranger dans un but expérimental précis. Des chromosomes humains artificiels ont même été fabriqués récemment et ils peuvent être introduits dans des cellules de mammifères, des cellules humaines ou d'une autre nature. La possibilité d'isoler un gène en tube à essai, de modifier spécifiquement sa structure et de le réinsérer ensuite dans l'organisme de départ a doté la dissection génétique d'un scalpel des plus précis.

L'utilisation de gènes spécifiques comme **marqueurs** est une autre des techniques utilisées avec succès. De

même qu'on utilise des étiquettes de couleurs vives pour identifier certains animaux ou certaines plantes au cours d'études biologiques, les généticiens se servent de formes de gènes faciles à détecter pour suivre la trace de structures — chromosomes, cellules, individus spécifiques. Cette technique a trouvé des applications dans toute la gamme des disciplines biologiques, de la médecine légale à la biologie cellulaire, de l'évolution à l'écologie. Par exemple, on isole actuellement des gènes responsables de maladies humaines grâce à la proximité sur le chromosome de séquences marqueurs sans aucune relation avec eux (c'est la technique du clonage positionnel). Avec la possibilité de déplacer des gènes d'un organisme à un autre, les généticiens ont remplacé des gènes d'origine par des **gènes « rapporteurs »**, dont les fonctions sont plus faciles à détecter et à suivre expérimentalement. (Un gène rapporteur est un type de gène marqueur qui permet de repérer une activité plutôt qu'une structure). Le gène de la luciférase de la luciole a été utilisé de cette façon ; le gène peut être inséré dans des chromosomes animaux ou végétaux de sorte que les cellules brillent aux stades du développement auxquels le gène d'origine était actif. Un gène de méduse codant une protéine vert fluorescent est également utilisé comme rapporteur. Les souris portant ce gène dégageant une lumière verte lorsqu'on les irradie avec des UV (figure 1-10). Tout compte fait, le génie génétique a révolutionné les sciences biologiques et aucun biologiste ne peut prétendre aujourd'hui ignorer cet outil analytique puissant.

MESSAGE

L'analyse génétique est maintenant une technique essentielle dans tous les domaines de la biologie.



Figure 1-10 Souris transgéniques contenant le gène de méduse qui code la protéine vert fluorescent, inséré dans leur chromosomes. (KYODO News International/AP.)

Le plus grand succès de la biologie est peut-être d'avoir montré exactement comment les gènes fonctionnent ou, en d'autres termes, comment l'information devient *forme*. C'est une histoire fascinante qui s'est développée avec une rapidité étonnante. Aujourd'hui, le codage et le flux de l'information génétique dans les cellules sont le fondement de la pensée biologique moderne et le point de départ de toute démarche expérimentale. Il vaut la peine de rappeler comment circule l'information génétique ou, ce qui revient au même, la façon dont les gènes agissent, ce qu'on a appelé le *nouveau paradigme de la biologie*. La figure 1-11 schématise l'action des gènes dans une cellule **eucaryote** type. Les Eucaryotes sont les organismes dont les cellules possèdent un noyau délimité par une membrane. Les animaux, les plantes et les champignons sont tous des Eucaryotes. À l'intérieur du noyau se trouvent les chromosomes, tandis qu'à l'extérieur de celui-ci s'étend un réseau complexe de structures membranaires, notamment le reticulum endoplasmique et l'appareil de Golgi, ainsi que des organites comme les mitochondries et les chloroplastes.

Dans le noyau eucaryote, certains gènes codant des protéines synthétisent leur produit protéique de façon plus ou moins continue, tandis que d'autres sont activés ou inactivés en fonction des besoins de la cellule ou de l'organisme. Le signal qui active un gène peut venir de l'extérieur de la cellule ; c'est le cas par exemple de substances telles que les hormones stéroïdes. Mais ce signal peut également provenir de l'intérieur même de la cellule, d'un gène *régulateur* spécial dont la fonction est d'induire ou d'arrêter le fonctionnement d'autres gènes. Les substances régulatrices se lient à une région particulière du gène et permettent la synthèse de transcrits de l'ADN du gène. Disséminés dans la région codant une protéine d'un gène eucaryote, se trouvent des segments qui ne sont pas destinés à être traduits en protéine. Les régions non codantes du gène, appelées introns, sont éliminées du transcrit initial et ce qui reste de la séquence d'ARN est l'**ARN messager** ou **ARNm**. Les molécules d'ARNm passent par les pores nucléaires et gagnent le cytoplasme où des organites cytoplasmiques appelés *ribosomes* se fixent à l'ARNm et traduisent en protéine l'information contenue dans la séquence d'ARNm. L'ARNm passe dans des ribosomes, qui catalysent l'assemblage du polypeptide, la chaîne d'acides aminés qui constituera la structure primaire de la protéine. Chaque acide aminé est apporté au ribosome par une molécule spécifique d'**ARN de transfert** (**ARNt**) qui se fixe sur une unité codante spécifique de l'ARNm. Les ARNt sont synthétisés à partir de gènes d'ARNt. Aucun ARNt n'est jamais traduit en protéine. Ils se recyclent constamment, apportant leur acide aminé spécifique aux ribosomes. Le ribosome est constitué d'un ensemble complexe de protéines et de plusieurs sortes d'ARN appelés **ARN ribosomiaux** (**ARNr**). Les gènes des ARNr sont situés dans une région particulière des chromosomes appelée *l'organisateur nucléolaire*. Comme l'ARNt, l'ARNr n'est jamais traduit en protéine.

MESSAGE

Les gènes codant des protéines fonctionnent en deux étapes de transfert d'information :



Chaque gène code une protéine distincte qui remplit une fonction spécifique dans la cellule (c'est le cas par exemple de la protéine symbolisée par un carré de la figure 1-11) ou dans d'autres parties de l'organisme si elle est exportée (la protéine symbolisée par un rond). La synthèse des protéines destinées à l'exportation (protéines sécrétées) a lieu à la surface du reticulum endoplasmique rugueux, un ensemble de grandes vésicules aplaties. La chaîne d'acides aminés terminée passe dans la lumière du reticulum endoplasmique, où elle se replie spontanément dans certains cas pour adopter sa forme protéique. Les protéines peuvent être modifiées à cette étape ; elles passent ensuite par les citernes de l'appareil de Golgi, puis dans des vésicules sécrétoires qui fusionnent avec la membrane cellulaire et libèrent leur contenu à l'extérieur de la cellule.

Les protéines destinées à exercer leur fonction dans le cytosol ainsi que la plupart des protéines qui œuvrent dans les mitochondries et les chloroplastes sont synthétisées dans le cytosol par des ribosomes qui ne sont pas liés aux membranes. Les enzymes de la glycolyse, par exemple, sont synthétisées de cette façon. La synthèse de toutes les protéines utilise le même mécanisme avec les mêmes types d'ARNt. La protéine achevée se détache du ribosome et se replie dans la conformation appropriée à l'intérieur du cytosol. Les protéines destinées aux organites sont identifiées spécialement pour qu'elles soient dirigées et insérées dans ces organites. Les mitochondries et les chloroplastes possèdent leurs propres petites molécules d'ADN circulaires. La synthèse des protéines codées par des gènes des ADN mitochondriaux ou chloroplastiques s'effectue dans des ribosomes localisés à l'intérieur de ces organites. Les protéines des mitochondries et des chloroplastes ont donc deux origines distinctes : elles sont codées par le noyau et importées dans l'organite ou codées par l'organite et synthétisées dans celui-ci.

Les Procaryotes, tels que les bactéries, sont des organismes dont les cellules ont une structure plus simple. Il n'existe dans ces cellules ni noyau ni structure délimitée par une membrane. La synthèse des protéines y est dans ses grandes lignes analogue à celle des Eucaryotes et utilise également des ARNm, des ARNt et des ribosomes mais présente certaines différences importantes. Les Procaryotes, par exemple, ne possèdent pas d'introns.

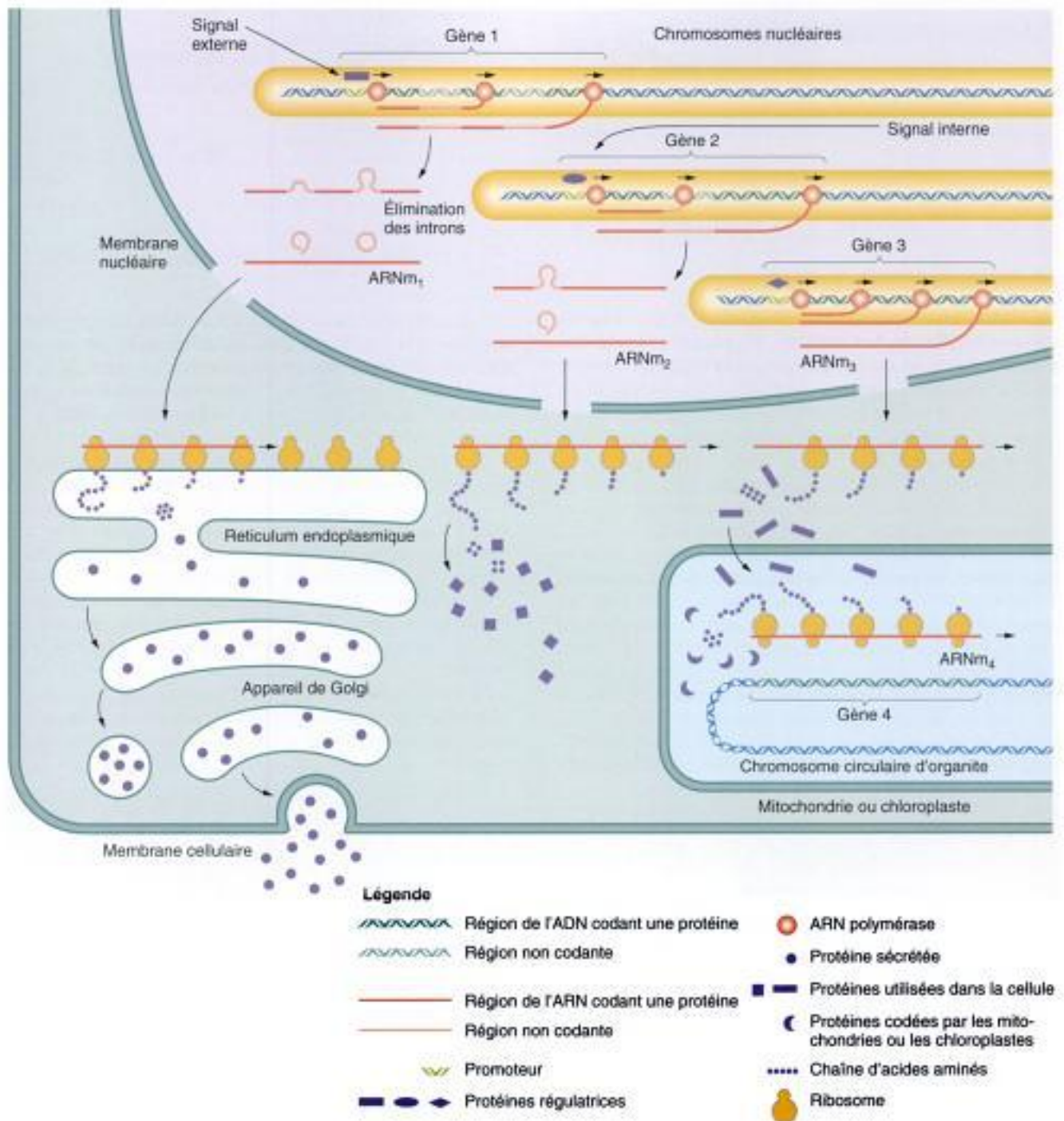


Figure 1-11 Vue simplifiée de l'action des gènes dans une cellule eucaryote. Le flux élémentaire d'information a lieu de l'ADN vers l'ARN puis vers les protéines. Quatre types de gènes sont présentés ici. Le gène 1 répond à des signaux régulateurs externes et fabrique une protéine destinée à être exportée. Le gène 2 répond à des signaux internes et code une protéine qui sera utilisée dans le cytoplasme. Le gène 3 code une protéine destinée à être transportée dans un organelle. Le gène 4 appartient à l'ADN d'un organelle et code une protéine destinée à être utilisée dans celui-ci. La plupart des gènes eucaryotes contiennent des introns, qui sont des régions (généralement non codantes) éliminées lors de la préparation de l'ARN messager fonctionnel.

Remarquez que de nombreux gènes d'organites possèdent des introns et qu'une enzyme de synthèse de l'ARN est nécessaire à la synthèse des ARNm des organites. Ces détails n'ont pas été représentés pour que la figure reste claire. (Les introns seront expliqués en détail dans les chapitres suivants.)

MESSAGE

Le flux d'information de l'ADN vers l'ARN, puis de l'ARN vers les protéines, constitue désormais un principe central de la biologie.

LA VARIATION EST À LA BASE DE LA GÉNÉTIQUE

Le concept de gène de Mendel était basé sur l'analyse de variants héréditaires — des organismes qui présentaient des différences identifiables et persistantes pour des caractères spécifiques. Mendel fut le premier à établir des règles pour l'analyse génétique de ces variants distincts. Ceci a constitué une étape importante car les questions concernant la variation sont prépondérantes dans le domaine de l'hérédité, et la plupart des analyses génétiques commencent avec des variants d'un certain type. Deux questions centrales de l'hérédité convergent vers la variation. Qu'est-ce qui détermine les caractéristiques fondamentales de différentes espèces (c'est-à-dire qu'est-ce qui est responsable de la variation *entre* espèces) ? Deuxièmement, qu'est-ce qui détermine la variation *à l'intérieur* d'une espèce ? Les gènes sont des éléments clés pour répondre à ces deux questions. Une espèce est unique en raison du groupe unique de gènes qu'elle comporte. Bien qu'on sache à présent qu'il existe des points communs considérables entre les génomes de différentes espèces, le génome de chaque espèce est néanmoins unique. L'un des buts de la génétique est de comprendre précisément la façon dont les gènes déterminent les caractéristiques

d'une espèce. Ce dessein est devenu bien plus accessible depuis les progrès réalisés dans la connaissance de la nature moléculaire des gènes et de leur fonctionnement.

La variation à l'intérieur d'une espèce est une observation que nous faisons chaque jour. C'est grâce à cette variation par exemple que nous distinguons les êtres humains les uns des autres et que nous reconnaissons nos animaux familiers. Cependant, la variation à l'intérieur d'une espèce peut avoir deux origines, la variation des gènes ou la variation de l'environnement. Dans chaque cas, il est donc nécessaire d'effectuer une analyse expérimentale pour mesurer l'importance relative des différents composants de la variation. Comment peut-il y avoir une variation dans l'ensemble des gènes caractéristique d'une espèce ? C'est-à-dire, si tous les membres d'une espèce ont le même ensemble de gènes, comment peut-il y avoir une variation génétique ? La réponse est que les gènes apparaissent sous différentes formes appelées **allèles**. Il peut par exemple, y avoir un seul gène responsable de la couleur d'une fleur, mais plusieurs allèles, chacun produisant une couleur de fleur différente. Dans une population, un gène peut avoir un allèle unique ou bien de nombreux allèles, mais puisque les organismes ne comportent qu'un ou deux jeux de chromosomes par cellule, chaque individu ne peut porter qu'un ou deux allèles d'un gène. On trouvera toujours les allèles d'un gène donné à la même position chromosomique. Les allèles d'un gène représentent différents niveaux d'activité, une nouvelle activité, ou aucune activité de ce gène. Un allèle est produit par mutation, ce qui implique un ou plusieurs changements de paires de nucléotides dans son ADN. Cette variation allélique est à la base de la variation héréditaire.

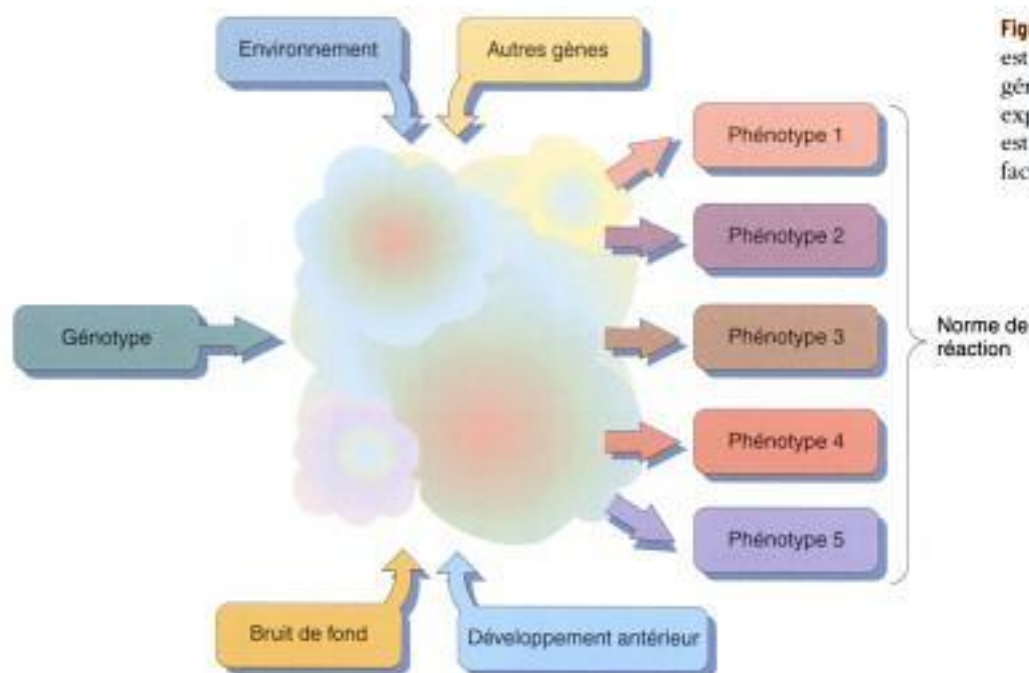
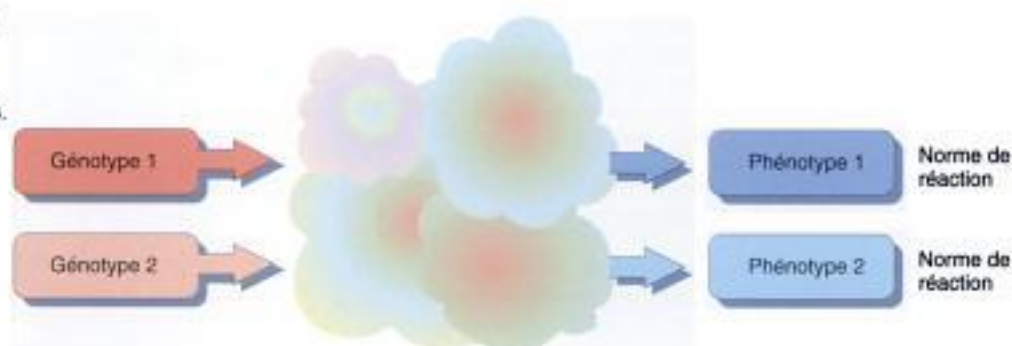


Figure 1-12 La norme de réaction est la gamme d'expression d'un génotype spécifique ; une expression variable de ce type est provoquée par certains des facteurs présentés ici.

Figure 1-13 Normes de réaction simples : des génotypes différents entraînent des phénotypes alternatifs distincts.



Il est important de comprendre les types de variation que l'on rencontre à l'intérieur d'une espèce. Un système de classification utile divise la variation en variation continue et discontinue. Dans le cas de la **variation discontinue**, on trouve un caractère dans la population sous deux formes distinctes ou plus. La forme exprimée d'un caractère s'appelle le **phénotype**. Par exemple, l'un des caractères étudiés par Mendel était la couleur des fleurs de pois. Il en existait deux phénotypes distincts, violet et blanc. Ces deux phénotypes étaient dus à deux allèles différents d'un même gène responsable de la couleur : l'allèle violet conférait la capacité de synthétiser le pigment violet et l'allèle blanc caractérisait l'absence de cette capacité, ce qui aboutissait à la blancheur de la fleur. On s'aperçoit souvent en réalité que des phénotypes contrastés sont déterminés par les allèles différents d'un même gène, comme le découvrit Mendel pour la couleur des pétales. L'ensemble spécifique des allèles portés par un individu est appelé le **génotype** ; c'est la base héréditaire du phénotype. Toutefois, l'ADN d'un gène ne peut produire spontanément un phénotype ; l'ADN dans un tube à essai est inerte. Pour

relier le génotype au phénotype, les normes de réaction des différents allèles ne se recouvrent pas. Ceci permet de prévoir une relation univoque entre le génotype et le phénotype dans la plupart des environnements et des conditions génétiques (figure 1-13). Dans l'exemple du travail de Mendel cité plus haut, l'allèle violet conduira presque toujours à des pétales violets et l'allèle blanc à des pétales blancs. Pour cette raison, la variation discontinue est utilisée avec succès par les généticiens pour identifier les allèles responsables d'un phénotype et leurs fonctions cellulaires. De plus, les variants discontinus dus à des allèles constituent le point de départ de la plupart des recherches génétiques.

Dans une population naturelle, l'existence de deux phénotypes alternatifs *courants* ou plus est appelée **polymorphisme** (en grec : « nombreuses formes ») ; un exemple est présenté dans la figure 1-14a. Les différents phénotypes sont parfois appelés **morphes**. Ces morphes sont souvent déterminés par les allèles d'un même gène. Pourquoi les populations présentent-elles un polymorphisme génétique ? Des types particuliers de sélection natu-

encore, dans de nombreux cas, les phénotypes sauvage et mutant sont déterminés par les allèles d'un même gène. Les mutants peuvent apparaître spontanément dans la nature ou être obtenus après traitement en laboratoire de populations par des agents mutagènes (produits chimiques ou radiations).

MESSAGE

On découvre souvent que les différentes formes phénotypiques d'un caractère au sein d'une variation discontinue sont déterminées par les allèles d'un même gène.

La **variation continue** d'un caractère présente souvent un éventail continu de phénotypes dans la population. Les caractères mesurables tels que la taille, le poids et l'intensité de la couleur sont de bons exemples de variation continue. Les phénotypes intermédiaires sont généralement plus courants que les phénotypes extrêmes, et lorsque les fréquences phénotypiques sont représentées sous la forme d'un graphe, on observe une distribution en cloche des phénotypes. (La variation continue est représentée et opposée à la variation discontinue dans une population naturelle dans la figure 1-15). Dans certaines distributions continues, toute la variation est d'origine environnementale et n'a aucune cause génétique. Dans d'autres cas, il y a une composante génétique due à une variation allélique d'un petit ou d'un grand nombre de gènes. Dans la plupart des cas, il y a à la fois une variation génétique et une variation due à l'environnement.

Dans les distributions continues, les normes de réaction des différents génotypes sont complexes et il n'y a pas de correspondance « terme à terme » entre le génotype et le phénotype (figure 1-16). Pour cette raison, on sait peu de choses sur les types de gènes responsables de variation continue. Depuis quelques temps cependant, on dispose de nouvelles techniques moléculaires plus puissantes qui permettent d'identifier et de caractériser ces gènes.

Dans la vie quotidienne, la variation continue est plus fréquente que la variation discontinue. Nous pouvons tous citer des cas de variation continue que nous avons observés dans des populations végétales ou animales (taille des fruits, rendement agricole, etc.) ; les populations humaines fourmillent elles aussi d'exemples. Un des domaines de la génétique où la variation continue est importante, est celui de la sélection des animaux et des végétaux. Un grand nombre des caractères visés par les programmes de sélection, tels que le poids des graines ou la production de lait, présentent une variation continue dans les populations. Les animaux ou les végétaux situés à une extrémité de la gamme sont choisis et croisés de façon sélective. Avant d'entreprendre une telle sélection, il est nécessaire d'effectuer des analyses pour évaluer la part respective des composants génétiques et environnementaux de la variation.

Certains phénotypes à variation *discontinue* présentent une **transmission complexe** qui ressemble à celle de la variation continue. Dans les populations humaines, les maladies cardiaques, les diabètes, les fentes palatines

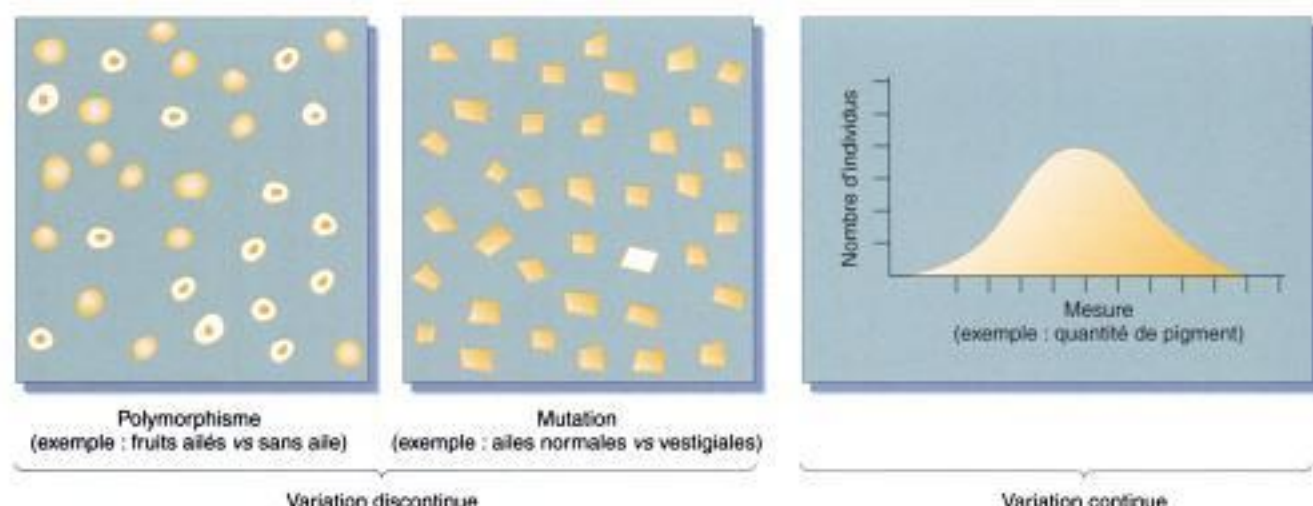
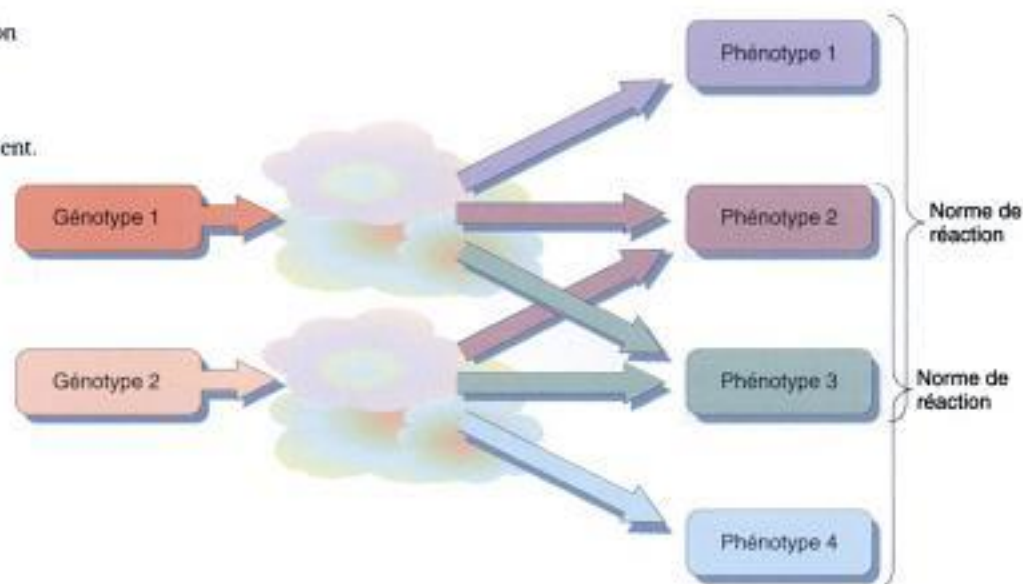


Figure 1-15 Variation discontinue et continue dans des populations naturelles. Dans les populations présentant une variation discontinue pour un caractère donné, les individus présentent tous un seul des différents phénotypes possibles. Dans le premier exemple, une population de plantes manifeste deux sortes de fruits courants différents, ailés et sans aile (voir figure 1-14a). On appelle ce genre de variation un *polymorphisme*. Parfois, la plus grande partie d'une population est d'un type, une mutation donnant lieu seulement à des alternatives occasionnelles, comme c'est le cas pour les drosophiles à ailes vestigiales. Les variants discontinus sont souvent déterminés par les allèles d'un même gène. La variation continue, quant à elle, ne présente pas d'alternatives discrètes de ce type : on peut trouver dans la population, un caractère exprimé suivant des degrés phénotypiques graduels. Il peut n'y avoir aucune cause génétique — toute la variation peut être due à l'environnement — ou des gènes peuvent être impliqués, souvent plusieurs, parfois un grand nombre.

Figure 1-16 Normes de réaction complexes. Des génotypes différents sont exprimés dans des gammes de phénotypes qui se chevauchent.



ou labiales et les sténoses du pylore (blocage de la sortie de l'estomac) en sont des exemples. Le risque d'une naissance avec ces affections est plus élevé dans les familles dont certains membres sont atteints ; il semble donc y avoir une composante génétique, mais la transmission ne suit pas le modèle allélique observé par Mendel. On pense que ces phénotypes reposent sur des interactions multiples entre les gènes et l'environnement, avec probablement un seuil physiologique au-delà duquel le phénotype variant est exprimé.

Nous reviendrons à la génétique quantitative dans les chapitres 12 et 18, mais la majorité du livre traitera de gènes responsables de variation discontinue simple, déterminée par les allèles de gènes uniques.

LA GÉNÉTIQUE EXPÉRIMENTALE

L'étude des gènes s'est avérée être une approche puissante pour comprendre les systèmes biologiques. Les gènes affectent quasiment chaque aspect de la structure et de la fonction d'un organisme, c'est pourquoi être capable d'identifier et de déterminer leur fonction est une étape importante pour élucider les différents processus qui sous-tendent un caractère particulier en cours d'étude. La génétique est une approche de l'ensemble de la biologie, et pas seulement de l'hérédité. Que signifie donc exactement l'approche génétique ? De nombreuses méthodologies sont utilisées pour étudier les gènes et leurs activités, mais on peut les résumer brièvement comme suit :

1. *Isolement de variants phénotypiques.* La plupart du temps, les variants utilisés sont des mutants. Chaque mutant représente potentiellement un composant génétique du système étudié.
2. *Analyse des descendants d'accouplements (- croisements -) contrôlés entre des variants phénotypiques.*

Ce type d'analyse, commencé par Mendel, identifie des gènes et leurs allèles, leurs positions sur les chromosomes, la façon dont ils sont transmis et leurs interactions.

3. *Analyse biochimique de processus cellulaires contrôlés par des gènes spécifiques.* Fondamentalement, la vie est un ensemble complexe de réactions chimiques. C'est pourquoi, étudier la façon dont les gènes contrôlent ces réactions est important pour disséquer cette chimie complexe. Les allèles entraînant une fonction biochimique déficiente sont inestimables dans ce type d'analyse.
4. *Analyse microscopique.* L'étude microscopique des mouvements et de la structure des chromosomes a longtemps été partie intégrante de l'analyse génétique. Grâce à de nouvelles technologies, on peut actuellement marquer les gènes eux-mêmes et leurs produits, ce qui permet, sous le microscope, de visualiser leur localisation.
5. *Analyse de l'ADN.* Plusieurs techniques, telles que le clonage, permettent d'isoler l'ADN d'un gène donné, pour l'analyser directement. Ceci permet des applications secondaires telles que la transgénèse, la manipulation des gènes et le séquençage de l'ADN.

Dans toute analyse génétique, ces méthodologies peuvent être utilisées pour étudier une fonction biologique donnée. Les approches sont très complémentaires et représentent un arsenal puissant et fin pour la recherche, ce dont nous verrons de nombreux exemples dans ce livre. La façon dont ces méthodologies ont été utilisées pour identifier la nature des gènes est présentée dans Les grandes découvertes de la génétique 1-2 à la page suivante.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 1-2

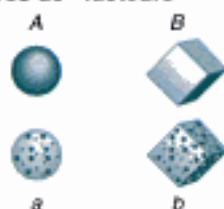
Déduire la nature des gènes

La compréhension actuelle de la structure et de la fonction des gènes est le résultat d'environ 100 ans de recherche. Il est donc important de retracer certaines des expériences décisives qui ont contribué à notre compréhension du gène en tant que concept, puis en tant que réalité. Nous reviendrons sur toutes les découvertes mentionnées ci-dessous tout au long de cet ouvrage : elles sont répertoriées ici dans l'ordre chronologique pour donner un aperçu de l'évolution de la génétique. Cet aperçu comporte également le nom des scientifiques à l'origine de ces découvertes, pour rappeler que ce sont des personnes qui font avancer la connaissance scientifique, grâce à leur génie et à leur dévouement. Une description plus détaillée de certaines découvertes historiques apparaîtra dans ce livre sous forme d'encadrés.

Bien qu'il soit probable que l'espèce humaine — et même ses ancêtres — ait toujours été intéressée par l'hérédité, les plus anciennes traces d'idées sur les mécanismes de l'hérédité remontent aux Grecs anciens. Ainsi, Aristote pensait que, puisque le sang irrigue et nourrit les organes du corps, la procréation devait elle aussi passer par l'intermédiaire du sang. Pour lui, la semence masculine était du sang purifié et la contribution génétique féminine à la génération suivante était le sang menstruel. Toutefois, ces idées et bien d'autres encore ne contribuèrent pas ou peu à la compréhension actuelle de l'hérédité. Après les Grecs, il y eut un silence de 2000 ans sur les mécanismes de l'hérédité. La génétique en tant que discipline scientifique n'existait pas avant les travaux de Gregor Mendel, au milieu du dix-neuvième siècle. Ceci s'explique par le fait que Mendel fut la première personne à concevoir l'idée d'un gène, et bien qu'à ce moment-là il ignorait tout de la nature physique ou chimique des gènes, ceux-ci furent à la base de la compréhension actuelle de l'hérédité. On peut faire remonter toute la recherche actuelle en génétique à la découverte de Mendel. Suivez la discussion sur l'évolution de la compréhension génétique en utilisant la figure qui l'accompagne et qui représente des schémas de certaines des étapes clés abordées.

Avant Mendel, la réflexion sur l'hérédité était basée principalement sur la logique et sur des conjectures et non sur l'expérimentation. C'est Mendel qui aborda la génétique comme une discipline expérimentale. Dans le jardin de son monastère, il réalisa pendant plusieurs années un grand nombre de pollinisations croisées expérimentales entre des plants de pois, qui présentaient une expression contrastée de certains caractères, tels que la couleur jaune ou verte des graines et la grande taille ou le nanisme des plants. Dans certains types de croisements, il observa des rapports mathématiques précis et reproductibles dans la descendance, et par de brillantes déductions à partir de ces rapports, il conclut non seulement à l'existence de gènes (qu'il appela « facteurs »), mais également à leur présence par paires dans les plants de pois et à la séparation de ces paires durant la formation des ovules et du pollen (voir partie 1 de la figure, montrant une paire de gènes A et a et une deuxième paire de

1. Paires de «facteurs»



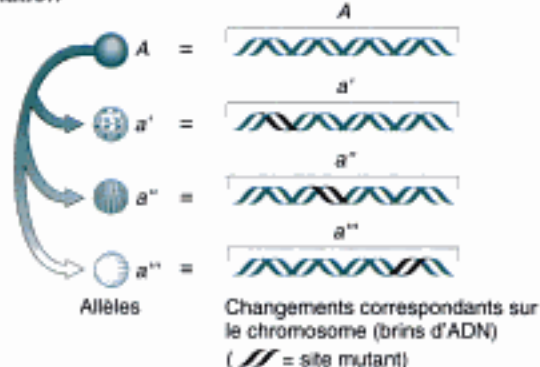
2. Gènes sur des chromosomes



3. Liaison génétique



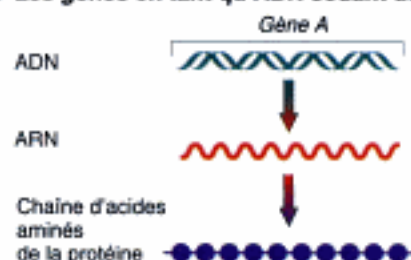
4. Mutation



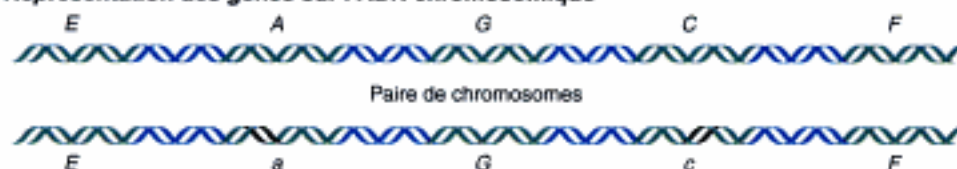
gènes B et b). Des propriétés différentes telles que le nanisme ou la grande taille des plants sont dues à des formes différentes d'un même type de gène (ce que nous appelons à présent des *allèles*). Mendel analysa également des plantes qui différaient par deux caractères. Il déduisit des rapports de la descendance de ces croisements, que des paires différentes de gènes sont réparties indépendamment dans les ovules et le pollen. Remarquez que dans cette figure, au-dessus des allèles sont indiquées des lettres symboles (par exemple, la paire A et a ; il s'agit d'une convention en génétique, comme nous le verrons plus loin).

Les gènes de Mendel n'étaient que des entités hypothétiques, des facteurs dont il supposait l'existence pour expliquer ses résultats. Le vingtième siècle fut le témoin de progrès considérables dans le développement de notre compréhension de la nature et du fonctionnement des gènes. Les publications de Mendel restèrent dans l'oubli jusqu'en 1900, jusqu'à ce que plusieurs autres chercheurs parvinrent indépendamment aux mêmes conclusions. Apparurent ensuite des centaines de publi-

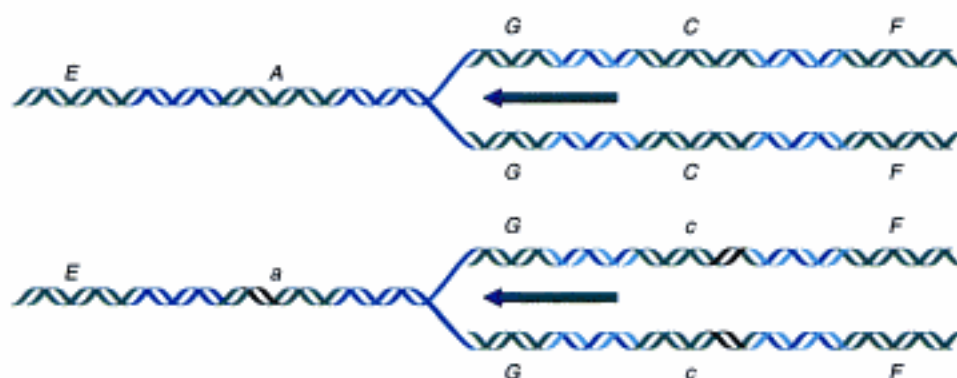
5. Les gènes en tant qu'ADN codant des protéines



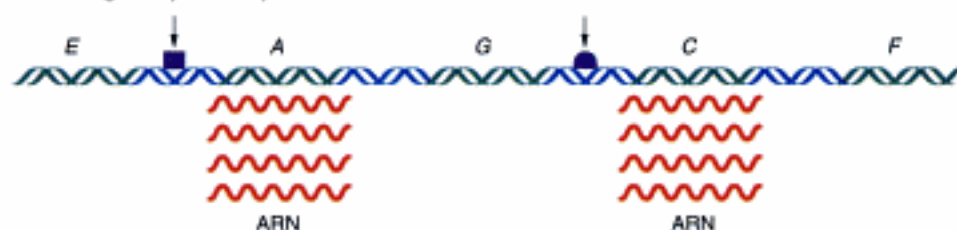
6. Représentation des gènes sur l'ADN chromosomique



7. Réplication semi-conservative des chromosomes



8. Gènes régulés par une protéine se fixant à l'ADN



cations faisant état d'une transmission « mendélienne » dans une vaste gamme de plantes et d'animaux, y compris les hommes. Il semblait alors que les idées de Mendel sur les gènes et leur profil de transmission fussent généralisables. Au début du vingtième siècle, les biologistes remarquèrent que la transmission des gènes suivait de près la transmission des chromosomes au cours des divisions nucléaires appelées méioses, qui se produisent pendant les divisions cellulaires, lors de la formation des ovules et des spermatozoïdes. Il semblait que les gènes fussent être des éléments des chromosomes (voir partie 2 de la figure), organisés par paires car chez la plupart des animaux et des végétaux, les chromosomes existent sous la forme de paires. Cette idée était soutenue par la démonstration de Thomas Hunt Morgan (1910) de la transmission parallèle de certains gènes de la mouche du vinaigre et des chromosomes

déterminant le sexe ; c'est-à-dire que le profil de transmission de certains gènes semblait suivre la transmission des chromosomes sexuels, suggérant une fois encore que les gènes se trouvaient sur les chromosomes. D'autres expériences réalisées par Morgan (1911) et l'un de ses étudiants, Alfred Sturtevant (1913), montrèrent également une transmission parallèle de certains gènes, suggérant que ces gènes se trouvaient sur le même chromosome — le concept de liaison génétique. Dans la partie 3 de la figure, par exemple, l'allèle A peut être transmis en même temps que l'allèle C plus souvent que le hasard ne le laisserait penser, suggérant que ces allèles sont liés sur le même chromosome. C'est le même cas pour les allèles a et c. Les chercheurs mirent au point une méthode pour déterminer la distance entre des gènes situés sur le même chromosome, en mesurant la fréquence à laquelle la transmission parallèle de

deux gènes est perturbée, débouchant sur de nouvelles combinaisons d'allèles (par exemple Ac et α Cl). On imagina que ces nouvelles combinaisons étaient provoquées par un processus de cassure et de réunion chromosomique contrôlé, appelé *crossing-over*. Plus deux gènes liés sont éloignés l'un de l'autre (c'est-à-dire plus leur distance génétique est importante), plus il est probable qu'il se produira un *crossing-over* entre eux, perturbant leur transmission concomitante. En 1916 un autre étudiant de Morgan, Calvin Bridges, prouva de façon indiscutable que les gènes sont situés sur les chromosomes. En utilisant des mouches du vinaigre avec un chromosome supplémentaire, il montra que la transmission anormale de certains gènes ne pouvait s'expliquer que s'ils appartenaient à ce chromosome supplémentaire. Finalement en 1931, Harriet Creighton et Barbara McClintock démontrèrent que la production de nouvelles combinaisons de gènes sur le même chromosome est en effet corrélée à un échange physique de certaines parties des chromosomes.

En 1927, Herman Muller montra que des mutations (partie 4) peuvent être produites à haute fréquence en traitant des cellules par des rayons X. Ceci constitua la première démonstration de l'existence d'un agent environnemental mutagène. (Les mutations peuvent également apparaître spontanément, mais à une fréquence plus faible). Le travail de Muller sur l'induction de mutations ouvrit la porte à l'utilisation de mutations pour disséquer des processus biologiques, une démarche encore largement utilisée aujourd'hui. En 1943, Salvador Luria et Max Delbrück utilisèrent des bactéries pour montrer que la mutation est un processus aléatoire qui peut se produire à tout moment dans n'importe quelle cellule.

La génétique moléculaire débuta en 1941, lorsque George Beadle et Edward Tatum utilisèrent le champignon *Neurospora crassa* pour montrer que certains gènes impliqués dans la chi-

mie de la cellule agissent en codant des protéines catalytiques appelées enzymes. Des études ultérieures réalisées sur d'autres organismes étendirent ce concept pour montrer que la plupart des gènes codent certains types de protéines, que ce soit des protéines qui jouent un rôle fonctionnel telles que les enzymes ou des protéines qui jouent un rôle structural dans les cellules. Peu après, Oswald Avery, C. M. McLeod et M. McCarthy (1944) découvrirent que la constitution génétique d'une cellule bactérienne peut être modifiée en ajoutant de l'ADN exogène (le premier exemple de transgénèse), montrant ainsi que les gènes bactériens sont constitués d'ADN, une découverte qui fut plus tard étendue à tous les organismes. Une étape décisive fut franchie en 1953, lorsque James Watson et Francis Crick proposèrent un modèle de double hélice pour la structure de l'ADN. Ce modèle montrait que l'ADN pouvait se répliquer par un déroulement progressif des deux brins entrelacés de la double hélice et par une nouvelle synthèse utilisant les brins exposés comme matrices. Ils proposèrent que chacun des deux brins entrelacés d'ADN était une chaîne de groupements chimiques appelés *nucléotides*, dont on connaissait quatre types. Les protéines étant des chaînes d'acides aminés, il fut suggéré qu'une séquence nucléotidique spécifique d'ADN (un gène) pouvait contenir l'information codée spécifiant la séquence d'acides aminés et donc, la structure de la protéine (partie 5).

En 1955, Seymour Benzer, approfondissant des études réalisées sur la mouche du vinaigre, montra que des allèles mutants de virus bactériens résultent de changements dans des régions localisées à l'intérieur du gène (sites mutants — se référer à la partie 4) et que les sites mutants à l'intérieur d'un gène peuvent être cartographiés les uns par rapport aux autres, ce qui permet d'obtenir une carte linéaire. Il était donc probable que le gène lui-même fut une structure linéaire comme dans le postulat de Watson et Crick. Il semblait également probable que l'ADN des

RÉSUMÉ

La génétique est l'étude des gènes à tous les niveaux, des molécules aux populations. La forme moderne de cette discipline apparut en 1860, avec les travaux de Gregor Mendel, qui fut le premier à suggérer l'idée de l'existence des gènes. On sait à présent qu'un gène est une région fonctionnelle appartenant à une longue molécule d'ADN qui constitue la structure fondamentale d'un chromosome.

La génétique a eu une influence profonde sur les activités humaines. La plupart de nos aliments et de nos vêtements proviennent d'organismes génétiquement améliorés, et le génie génétique moléculaire est en train d'étendre le champ de la génétique appliquée. De plus, une fraction élevée des maladies humaines comportent une composante génétique. La recherche en génétique a également changé la façon dont nous nous percevons par rapport au monde organique et au reste de l'univers.

La génétique a révolutionné la biologie en montrant que la plupart des organismes de la planète fonctionnent à partir d'un système commun de stockage et d'expression de l'information, centré sur l'ADN — un système dans lequel l'information circule de l'ADN vers l'ARN puis vers les pro-

téines. De plus, la dissection génétique est un outil pointu d'analyse, utilisé dans toutes les sciences biologiques.

Les gènes n'agissent pas de manière isolée. Ils interagissent au contraire à de nombreux niveaux ensemble et avec l'environnement en produisant un phénotype. La relation du génotype avec le phénotype selon l'environnement dans lequel il s'exprime est appelé la *norme de réaction*. La plupart des progrès en génétique sont venus d'études en laboratoire, portant sur des caractères ayant une correspondance simple, terme à terme entre le génotype et le phénotype. Cependant, dans les populations naturelles, la variation phénotypique présente généralement une relation plus complexe avec le génotype et non une correspondance terme à terme. Les variants discontinus ont été le point de départ de la plupart des expériences en analyse génétique. Souvent, les phénotypes distincts d'un même caractère sont déterminés par les allèles d'un même gène.

Les principaux outils de la génétique sont l'analyse des croisements de variants, la biochimie, la microscopie et le clonage de l'ADN.

gènes appartenait à une molécule continue d'ADN, le chromosome (partie 6). Ceci fut d'ailleurs prouvé plus tard. Le processus de séparation des brins pour la réplication de l'ADN (appelé processus semi-conservatif — partie 7) fut démontré expérimentalement pour la première fois par Matthew Meselson et Frank Stahl en 1958. À la fin de la réplication d'une molécule d'ADN, il y a deux molécules. Chacune d'elles est un hybride constitué d'un brin parental entrelacé avec un brin néo-synthétisé, d'où le terme de réplication semi-conservative. En 1961, Francis Crick et Sydney Brenner montrèrent que l'information codée dans la séquence nucléotidique devait être traduite par groupes de trois nucléotides pour chaque acide aminé. Charles Yanofsky (1966) montra que les positions des sites mutants à l'intérieur d'un gène correspondent parfaitement aux positions des acides aminés modifiés dans la séquence polypeptidique de la protéine correspondante. Le « dictionnaire » génétique complet des 64 triplets de codage (codons) et les acides aminés auxquels ils correspondent, fut déduit en 1966 par Marshall Nirenberg et Gobind Khorana. Des études ultérieures réalisées sur d'autres organismes montrèrent que la structure en double hélice de l'ADN, son mode de réplication et le dictionnaire des codons sont les mêmes chez quasiment tous les organismes, que ce soit des plantes, des animaux, des champignons ou des bactéries.

En 1961, François Jacob et Jacques Monod établirent le modèle de la régulation des gènes en montrant que des gènes bactériens pouvaient être activés (pour produire des transcrits d'ARN) et réprimés par la fixation de protéines régulatrices à une région adjacente au site de début de la synthèse d'ARN (partie 8).

Les progrès techniques ont joué un rôle important dans l'avancée de la connaissance de la génétique. Daniel Nathans et Hamilton Smith découvrirent en 1970 une catégorie particulière d'enzymes appelées enzymes de restriction qui coupaient l'ADN au niveau de séquences cibles spécifiques. Cette découverte

permit à Paul Berg en 1972 de fabriquer la première molécule d'ADN recombinant, composée d'ADN issu de deux organismes différents. Il isola des molécules de deux origines distinctes, les coupa et les relier l'une à l'autre dans un tube à essai. Ces progrès permirent en 1973 à Herbert Boyer et Stanley Cohen d'isoler des gènes individuels et de les amplifier en un nombre important de copies. Ils insérèrent ensuite les gènes dans des molécules d'ADN auto-réplicantes (de petits chromosomes de différents types) et les introduisirent dans des cellules vivantes. De cette pratique naquit la technologie de l'ADN recombinant qui domine actuellement la génétique. En 1977, deux techniques différentes furent inventées pour déterminer la séquence nucléotidique de l'ADN, l'une par Allan Maxam et Walter Gilbert et l'autre par Fred Sanger. Il fut alors possible d'examiner la structure des gènes directement par séquençage nucléotidique, confirmant donc des déductions sur les gènes qui n'avaient jusque-là été faites qu'indirectement. En 1985, Kary Mullis inventa une technique appelée réaction en chaîne de la polymérase (PCR) qui permet de trouver et d'amplifier des segments spécifiques d'ADN à partir de mélanges complexes. En 1985, Michael Smith inventa la mutagenèse dirigée, une technique qui permet d'introduire des changements nucléotidiques spécifiques dans un gène choisi.

Les années 1980 et 1990 virent le début de projets visant à séquencer intégralement les génomes de plusieurs organismes importants en recherche génétique. C'est en 1990 que le projet de séquençage du génome humain débuta. En 1995, le premier séquençage complet du génome d'un organisme fut annoncé. Il s'agissait de celui d'une bactérie, *Haemophilus influenzae*. Le premier séquençage intégral d'un génome eucaryote fut celui de la levure de boulanger, *Saccharomyces cerevisiae*, publié en 1996. Le séquençage de plusieurs autres génomes eucaryotes est presque terminé et la fin du séquençage humain est prévue pour 2005.

DIAGRAMME CONCEPTUEL

Chaque chapitre comporte un exercice qui consiste à établir un diagramme conceptuel. En établissant ces diagrammes, vous organiserez les connaissances que vous venez juste d'acquérir. En apprenant, nous devons intégrer nos nouvelles connaissances dans le cadre de la discipline. Une partie essentielle de cette tâche consiste à discerner les liens qui existent entre de nouvelles connaissances et le savoir acquis précédemment. Ce n'est pas aussi facile que cela paraît et l'esprit peut parfois nous jouer des tours en nous faisant croire que nous avons une vision d'ensemble du sujet. Les diagrammes conceptuels sont un bon moyen de se prouver que nous savons véritablement de quelle façon les différents processus, structures et idées de la génétique sont reliés. Ces diagrammes peuvent également nous aider à localiser les défauts de notre compréhension.

Une sélection de termes est indiquée et le défi consiste à tracer des lignes ou des flèches entre les termes que vous pensez reliés, en inscrivant sur la flèche une description précise de la relation qui existe entre ces termes. Tracez autant de relations que possible. Forcez-vous à établir des connexions, sans tenir compte de l'éloignement des deux termes au premier abord. La descrip-

tion au-dessus de la flèche doit être suffisamment précise pour qu'une autre personne, à la lecture de votre diagramme, puisse comprendre à quoi vous pensez. Parfois le diagramme révèle une simple succession d'étapes consécutives, parfois un réseau complexe d'interactions. Il n'y a pas de bonne réponse ; il n'y a que de nombreuses variations, toutes correctes. Il peut bien sûr vous arriver d'établir quelques interconnexions inexactes, et découvrir ces confusions est également un des buts de l'exercice.

La figure 1-17 à la page suivante présente un exemple de diagramme conceptuel mettant en relation les termes suivants, issus de différents chapitres :

substitution de paire de bases / récessif / rapport 3 : 1 / albinos / environnement / mélanine / voie métabolique / site actif / enzyme / gène.

À présent, établissez un diagramme conceptuel pour ce chapitre, en reliant le maximum des termes suivants. Remarquez que ces termes ne sont pas indiqués dans un ordre particulier.

génotype / phénotype / norme de réaction / environnement / développement / organisme.

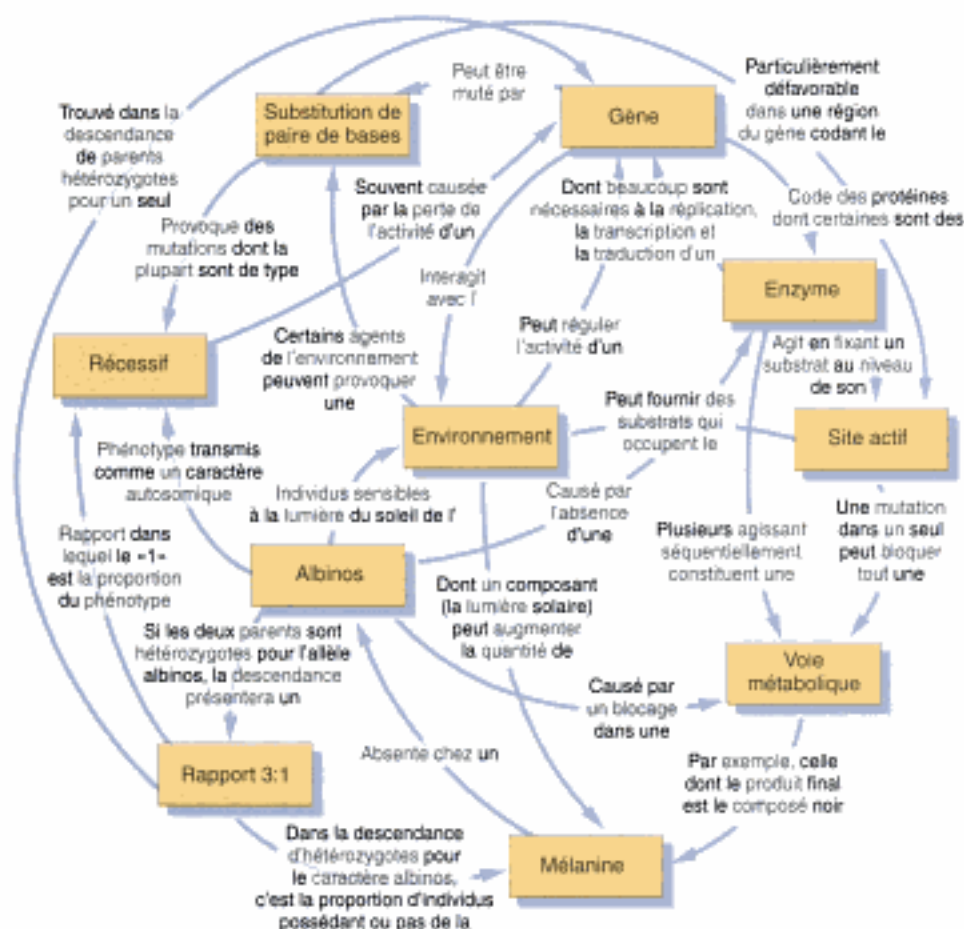


Figure 1-17 Exemple de diagramme conceptuel.

QUESTIONS GÉNÉRALES

1. Définissez la *génétique*. Pensez-vous que les éleveurs de chevaux de course de l'Égypte ancienne étaient des généticiens ? À votre avis, en quoi leurs méthodes différaient-elles de celles des généticiens actuels ?
2. Citez trois manières par lesquelles la génétique a affecté la société moderne.
3. Comment l'ADN dicte-t-il les propriétés générales d'une espèce ?
4. Quelles sont les deux caractéristiques de l'ADN qui lui permettent de tenir son rôle de molécule héréditaire ? Pouvez-vous imaginer d'autres types de molécules héréditaires que l'on pourrait trouver chez des formes de vies extraterrestres ?
5. Quelle est la relation entre les normes de réaction et la variation phénotypique à l'intérieur d'une espèce ?
6. Quels sont les types et l'importance de la variation phénotypique dans une espèce ?
7. La formule « *génotype* + *environnement* = *phénotype* » est-elle juste ?
8. Citez cinq dilemmes d'ordre éthique en génétique.
9. Quel est le but de la dissection génétique ?
10. En quoi la génétique a-t-elle unifié la biologie ?

2

LA STRUCTURE DES GÈNES ET DES GÉNOMES



POUR L'ESSENTIEL

- L'ADN est une double hélice constituée de deux chaînes nucléotidiques entrelacées, antiparallèles et complémentaires.
- Un génome est composé d'une ou plusieurs molécules d'ADN ; chacune est organisée sous forme de chromosome.
- Les génomes procaryotes sont la plupart du temps des chromosomes circulaires uniques.
- Les génomes eucaryotes ont des composants situés dans leur noyau et d'autres dans leurs organites.
- La garniture nucléaire du génome est constituée d'un ou deux jeux de chromosomes linéaires.
- Dans un chromosome nucléaire, l'ADN est enroulé en spirale autour de protéines, les histones, ce qui conduit à une structure en solénoïde linéaire compacte.
- Un gène est une région chromosomique capable d'engendrer un transcrit fonctionnel.
- Les gènes de nombreuses espèces eucaryotes contiennent des régions non codantes appelées *introns*.

Chromosomes de fibroblastes de femelle de Muntjac, une petite biche d'Asie du Sud-Est et d'Inde ($2n=6$). Les six chromosomes visibles proviennent d'une cellule interrompue dans son processus de division mitotique. Les trois sortes de chromosomes ont été colorées à l'aide de sondes d'ADN spécifiques de chaque chromosome, chaque sonde étant associée à un colorant fluorescent distinct (« peinture chromosomique »). Un noyau provenant d'une autre cellule est figé entre deux divisions mitotiques : remarquez que les chromosomes sont dans un état plus étiré et semblent « remplir » tout le noyau.

(Photographie communiquée par Fergtang Yang et Malcolm Ferguson-Smith de l'Université de Cambridge. Couverture de *Chromosome Research* vol.6, No.3, Avril 1998.)

L'unité la plus petite de la vie est la cellule. Les organismes peuvent être unicellulaires, comme les bactéries, de nombreux champignons et algues, ou pluricellulaires, comme les végétaux et les animaux. Toutes les caractéristiques d'un organisme sont déterminées par la structure et la fonction des cellules qui le composent. Pour comprendre la façon dont les gènes déterminent les caractéristiques fondamentales d'une espèce et contribuent à la variation à l'intérieur de celle-ci, il est nécessaire de replacer les gènes et leur action dans le contexte cellulaire. En d'autres termes, il nous faut connaître la nature du matériel génétique, sa localisation à l'intérieur des cellules et la façon dont il remplit sa fonction dans le cadre cellulaire. Ce chapitre sert d'introduction à la nature et l'organisation cellulaire du matériel génétique chez plusieurs catégories d'organismes. Le chapitre 3 présente certains détails du fonctionnement du matériel génétique.

Comment l'ADN est-il organisé dans une cellule ? Le jeu complet d'ADN d'un organisme est appelé un **génome**. Un génome est composé de longues molécules d'ADN, qui sont à leur tour les principaux composants des **chromosomes**. Chaque chromosome contient une molécule d'ADN portant de nombreux gènes. Comme nous l'avons vu au chapitre 1, les gènes sont des régions de l'ADN chromosomique qui peuvent être transcrites en ARN. Les génomes de la plupart des organismes procaryotes sont constitués d'un seul chromosome, tandis que les génomes des Eucaryotes comportent plusieurs chromosomes. Chez les Eucaryotes, la plupart des chromosomes sont situés dans le noyau, mais les mitochondries et les chloroplastes contiennent chacun un type unique de chromosome.

Nous devons également prendre en compte les molécules d'ADN qui ne font pas partie du génome des organismes. Les cellules de nombreux Procaryotes et de quelques Eucaryotes contiennent des molécules d'ADN accessoires appelées **plasmides**. Ces molécules d'ADN plasmidique contiennent des gènes nécessaires à leur propre propagation, on les appelle donc (parfois de façon abusive) génomes plasmidiques. Les virus, qui ne sont pas des organismes, possèdent néanmoins leur propre jeu de gènes, qui est également désigné par le terme de **génome**.

Puisque les génomes de tous les organismes, de tous les plasmides et de la plupart des virus sont constitués d'ADN, notre discussion commencera par la structure de cette molécule qui est responsable des propriétés inhérentes au monde vivant.

LA NATURE DE L'ADN

Comment savons-nous que les génomes sont composés d'ADN ? En utilisant des techniques physiques et histo-chimiques, il est relativement facile de le démontrer pour les chromosomes nucléaires eucaryotes. Les colorants qui se fixent à l'ADN, tels que le Feulgen ou le DAPI marquent

avant tout les chromosomes nucléaires dans les cellules et dans une moindre mesure, les mitochondries et les chloroplastes. De plus, si une masse de cellules est broyée et que ses composants sont fractionnés, il devient évident que la majeure partie de l'ADN peut être isolée à partir de la fraction nucléaire et le reste à partir des mitochondries et des chloroplastes.

Le fait que l'ADN soit le matériel héréditaire a maintenant été prouvé pour de nombreux Procaryotes et Eucaryotes. Si des cellules d'un génotype (le receveur) sont mises en contact avec de l'ADN provenant d'un autre génotype (le donneur), l'ADN donneur peut être absorbé par les cellules receveuses. Parfois, un morceau d'ADN s'intègre dans le génome du receveur et modifie un aspect du phénotype du receveur en celui du donneur d'ADN. Un tel résultat démontre que l'ADN est la substance qui détermine le génotype et donc qu'il s'agit du matériel héréditaire (voir Les grandes découvertes de la génétique 2-1).

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 2-1

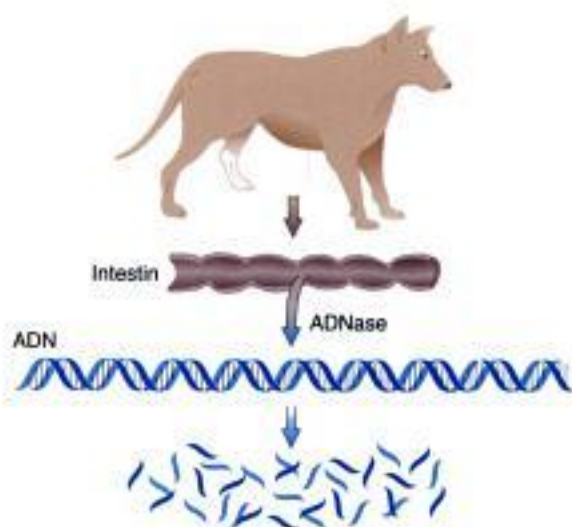
Oswald Avery démontre que le matériel héréditaire est l'ADN

En 1928, Frederick Griffith réussit à transformer de façon définitive une souche non virulente et non encapsulée de la bactérie *Pneumococcus*, en une souche virulente et encapsulée, en ajoutant un extrait de cellules mortes de ce dernier type. L'étape suivante consistait à découvrir le composant des cellules mortes qui était le « principe transformant », car cette substance conférait des propriétés héréditaires à la souche bactérienne receveuse.



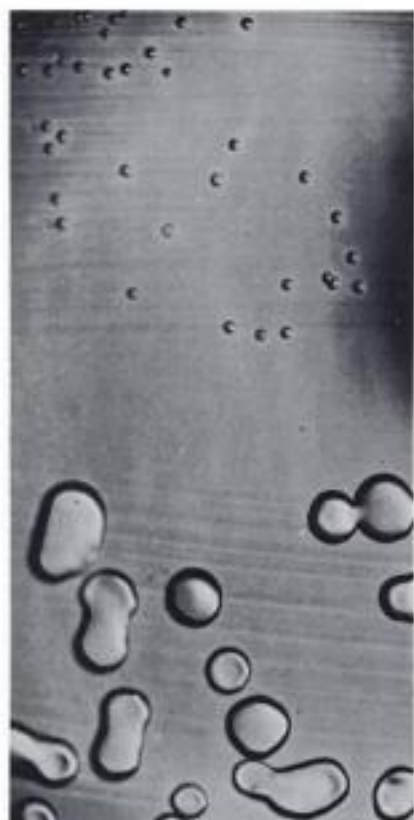
David M. Phillips/Visuals Unlimited.

En 1944, l'approche d'Oswald Avery consistait à inactiver une à une les principales catégories de substances chimiques contenues dans l'extrait et à vérifier si la propriété de transformation était elle aussi inactivée. Par exemple, l'un des candidats au titre de principe transformant était le matériel polysaccharidique de la capsid, qui forme une paroi lisse autour des cellules bactériennes. Pourtant, lorsque ce matériel était inactivé, la capacité de transformation du mélange persistait. Avery montra également que les protéines, les graisses et les acides ribonucléiques (ARN) n'étaient pas non plus le principe transformant. Au contraire, lorsque l'extrait était traité par l'enzyme désoxyribonucléotide polymérase, que l'on appelle à présent ADNase (extraite de muqueuses intestinales de chien ou de sérum de lapin ou de chien), qui coupe l'ADN, le principe transformant était détruit. Lorsque ces extraits de mucus ou de sérum étaient chauffés pour dénaturer l'ADNase, le principe transformant n'était pas affecté.



Dans une lettre de 1943 à son frère Roy, Avery écrit :

Durant les deux dernières années, d'abord avec MacLeod et à présent avec le Dr McCarthy, j'ai essayé de déterminer la nature chimique de la substance contenue dans l'extrait bactérien qui induit ce changement spécifique... Un travail accompagné de maux de tête et de déceptions. Je crois cependant que nous y sommes... Mais aujourd'hui, cela demande beaucoup de preuves argumentées pour convaincre tout le monde que le sel d'acide désoxyribonucléique, sans protéine, pourrait être doté de ces propriétés biologiques actives et spécifiques, et c'est ce que nous essayons de prouver. Il est très amusant d'échafauder des hypothèses, mais il est plus sage de les éprouver soi-même avant que quelqu'un d'autre ne s'y essaie.



D'après O. T. Avery, C. M. MacLeod et M. McCarty, *J. Exp. Med.* **79** (1944) : 158.

En ajoutant de l'ADN seul, Avery pouvait modifier le génotype et le phénotype de cellules bactériennes, démontrant ainsi de façon convaincante que l'ADN est le matériel génétique. On sait maintenant que les fragments de l'ADN transformant conférant la virulence pénètrent dans le chromosome bactérien et y remplacent leurs homologues responsables de l'avorulence.

Les trois rôles de l'ADN

Même avant la découverte de la structure de l'ADN, les études génétiques avaient clairement indiqué plusieurs propriétés que devait absolument présenter le matériel héréditaire.

Toutes les cellules du corps doivent avant tout avoir la même constitution génétique ; le matériel génétique doit donc être répliqué à l'identique à chaque division. Nous reviendrons plus loin dans ce chapitre sur les caractéristiques structurales de l'ADN qui permettent une duplication aussi fidèle.

Deuxièmement, le matériel génétique doit contenir des informations, car il doit coder la constellation de

protéines exprimées par un organisme. La façon dont l'information codée est déchiffrée en protéines fera l'objet du chapitre 3.

Enfin, même si la structure de l'ADN doit être relativement stable pour que les organismes puissent s'appuyer sur l'information qu'il code, il doit également permettre à l'information codée de changer en de rares occasions. Ces changements, appelés *mutations*, constituent le matériel de base — la variation génétique — sur lequel agit la sélection de l'évolution. Nous examinerons les mécanismes de mutation au chapitre 7.

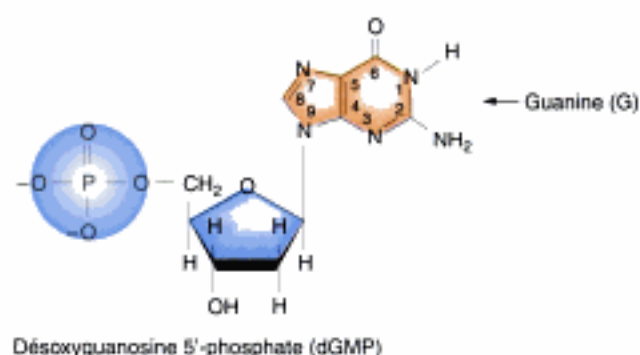
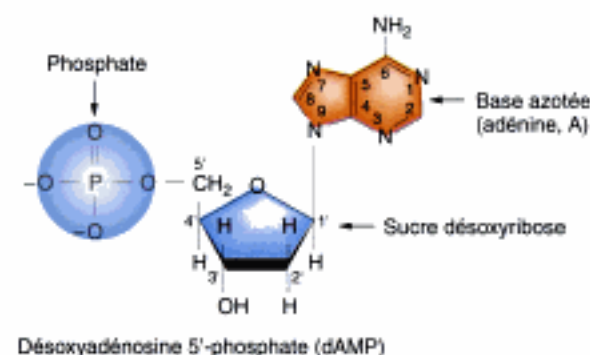
Les éléments de construction de l'ADN

L'ADN comporte trois types de composants chimiques : du **phosphate**, un sucre appelé **désoxyribose** et quatre **bases** azotées — l'adénine, la guanine, la thymine et la

cytosine. Deux de ces bases, l'adénine et la guanine possèdent une structure à deux cycles, caractéristique d'une substance chimique appelée *purine*. Les deux autres bases, la cytosine et la thymine, ont une structure à un seul cycle, d'un type appelé *pyrimidine*. Les composants chimiques de l'ADN sont organisés en groupes appelés **nucléotides**, composés chacun d'un groupement phosphate, d'une molécule de désoxyribose et de l'une des quatre bases. Il est pratique de faire référence à chaque nucléotide par la première lettre de son nom : A, G, C et T. La figure 2-1 présente la structure des quatre nucléotides de l'ADN.

Comment une molécule avec si peu de composants peut-elle remplir le rôle de molécule héréditaire ? Certains éléments de réponse furent apportés en 1953, lorsque James Watson et Francis Crick montrèrent de façon précise l'organisation des nucléotides de l'ADN (voir Les grandes découvertes de la génétique 2-2). La structure de l'ADN est résumée dans le paragraphe suivant.

Nucléotides puriques



Nucléotides pyrimidiques

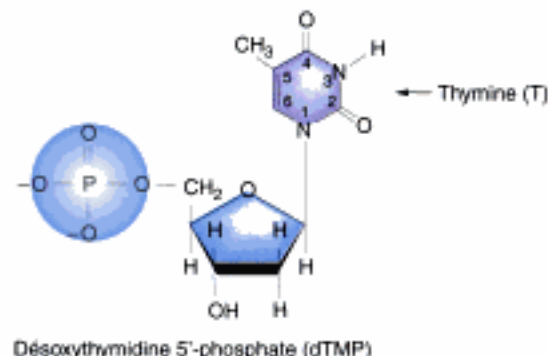
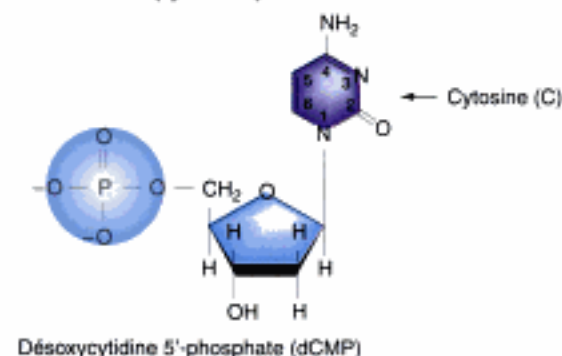


Figure 2-1 Structure chimique des quatre nucléotides (deux avec des bases puriques et deux avec des bases pyrimidiques) qui sont les éléments de construction de l'ADN. Le sucre est appelé désoxyribose car il s'agit d'une variation d'un sucre courant, le ribose, qui possède un atome d'oxygène de plus.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 2-2

James Watson et Francis Crick proposent la structure correcte de l'ADN

Un article de 1953 écrit par James Watson et Francis Crick dans la revue *Nature* commence par les deux phrases qui ouvrirent une nouvelle ère dans la biologie : « Nous aimerions suggérer une structure pour le sel d'acide désoxyribonucléique (ADN). Cette structure possède des caractéristiques nouvelles qui présentent un intérêt biologique considérable ». La structure de l'ADN était au centre d'un débat animé depuis 1944, lorsque Avery et coll. montrèrent que l'ADN est la substance héréditaire. Bien que la composition générale de l'ADN fût connue, on ignorait comment s'agençaient ses constituants. La structure devait rendre compte des deux propriétés indispensables à une molécule

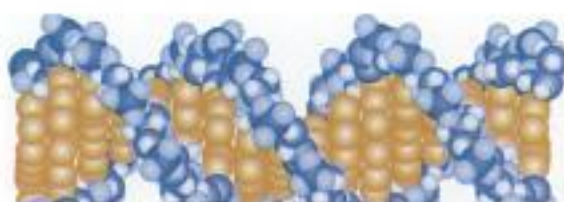


Motif de diffraction des rayons X.
Rosalind Franklin/Science Source/ Photo Researchers.

héréditaire, la capacité de stocker de l'information et la capacité de se répliquer.

Watson (un généticien américain travaillant sur les phages) était invité au laboratoire de Cavendish, à Cambridge, en Angleterre, lorsqu'il commença sa collaboration avec Crick, un physicien anglais. Leur modèle d'ADN était basé sur les données dont on disposait à l'époque sur la structure de l'ADN. Maurice Wilkins et Rosalind Franklin avaient obtenu des diagrammes de diffraction des rayons X qui suggéraient une structure hélicoïdale pour l'ADN. Erwin Chargaff avait analysé de nombreux ADN différents et en avait déduit des règles empiriques sur les quatre bases constituantes. La plus importante était une quantité égale d'adénine (A) et de thymine (T) d'une part et de guanine (G) et de cytosine (C) d'autre part.

À partir de ces informations, Watson et Crick rassemblèrent une structure qui semblait cohérente avec les différentes données. La structure qu'ils proposaient était une double hélice constituée de deux chaînes nucléotidiques orientées en sens inverse l'une par rapport à l'autre. Les deux chaînes étaient reliées l'une à l'autre par l'appariement de bases spécifiques, A avec T et G avec C. Ceci coïncidait avec l'égalité de A et T, et de C et G montrée par Chargaff. Pour finir, ils vérifièrent leurs idées en construi-



sant une maquette d'ADN en métal. Watson décrit plus tard en ces termes l'excitation ressentie en assemblant le modèle :

Je descendis pour voir si l'atelier pouvait aller plus vite et produire les bases puriques et pyrimidiques à la fin de l'après-midi. Seul un petit encouragement fut nécessaire pour que la tâche soit accomplie dans les heures qui suivirent. Les pièces de métal brillantes furent ensuite immédiatement utilisées pour fabriquer un modèle dans lequel pour la première fois, tous les composants de l'ADN allaient être présents. En une heure environ j'avais arrangé les atomes suivant des positions qui satisfaisaient à la fois les données obtenues par les rayons X et les lois de la stéréochimie. L'hélice résultante était de pas droit, avec ses deux chaînes orientées en sens inverse. Un modèle ne peut être manipulé facilement que par une seule personne. Aussi, Francis n'essaya pas de vérifier mon travail avant que je ne revienne et que je n'aie dit qu'à mon sens tout s'accordait. Un contact interatomique était un peu plus court que sa distance optimale ; il n'était cependant pas en désaccord avec plusieurs valeurs publiées et cela ne me gêna pas. Pendant quinze minutes encore Francis s'affaira sans rien trouver d'erroné. À de brefs intervalles, mon estomac se nouait lorsque je le voyais froncer les sourcils. Dans chaque cas, il finissait par être satisfait et se déplaçait pour vérifier qu'un autre contact interatomique était acceptable. Tout nous semblait donc correct lorsque nous sortîmes pour aller dîner.

La double hélice remplissait les exigences d'une substance héréditaire : la séquence nucléotidique pouvait coder la séquence d'acides aminés d'une protéine, et la réplication était possible par la séparation des brins et une nouvelle synthèse dictée par la spécificité de l'appariement des bases.



James Watson et Francis Crick devant leur modèle d'ADN.
Camera Press.

L'ADN est une double hélice

L'ADN est composé de deux chaînes (brins) appariées de nucléotides, enroulées en une double hélice. Les deux brins nucléotidiques sont maintenus ensemble par de faibles associations entre les bases de chaque brin, formant une structure semblable à un escalier en colimaçon (figure 2-2). Le squelette de chaque brin est un polymère sucré répété de phosphate et de désoxyribose. Les liaisons sucre-phosphate du squelette sont appelées **liaisons phosphodiester**. La fixation des liaisons phosphodiester aux groupements sucrés est importante pour décrire la façon dont chaque nucléotide est organisé. Remarquez que les carbones des groupements sucrés sont numérotés de 1' à 5'. Une partie de la liaison phosphodiester est établie entre le phosphate et le carbone 5' du désoxyribose, tandis que l'autre se trouve entre le phosphate et le carbone 3' du désoxyribose. On dit donc de chaque squelette

sucré-phosphate qu'il a une polarité 5'-3'. Comprendre cette polarité est essentiel pour comprendre la façon dont l'ADN remplit ses différents rôles. Dans la molécule d'ADN double-brin, les deux squelettes sont en sens opposé ou **antiparallèles**, comme le montre la figure 2-2. L'un des brins est orienté de 5' en 3' (5'→3'), de gauche à droite. L'autre brin est orienté de 5' en 3' de droite à gauche, autrement dit de 3' en 5' de gauche à droite (3'→5').

Les bases sont fixées au carbone 1' de chaque désoxyribose dans le squelette de chaque brin. Les interactions entre des paires de bases, une sur chaque brin, maintiennent les deux brins de la molécule d'ADN ensemble. Les bases de l'ADN interagissent selon une règle simple : il existe seulement deux paires de bases : A•T et G•C. Les bases de ces deux paires sont dites **complémentaires** l'une de l'autre. Ceci signifie que sur n'importe quelle « marche » de l'escalier d'ADN en colimaçon, les seules associations de bases qui puissent exister entre

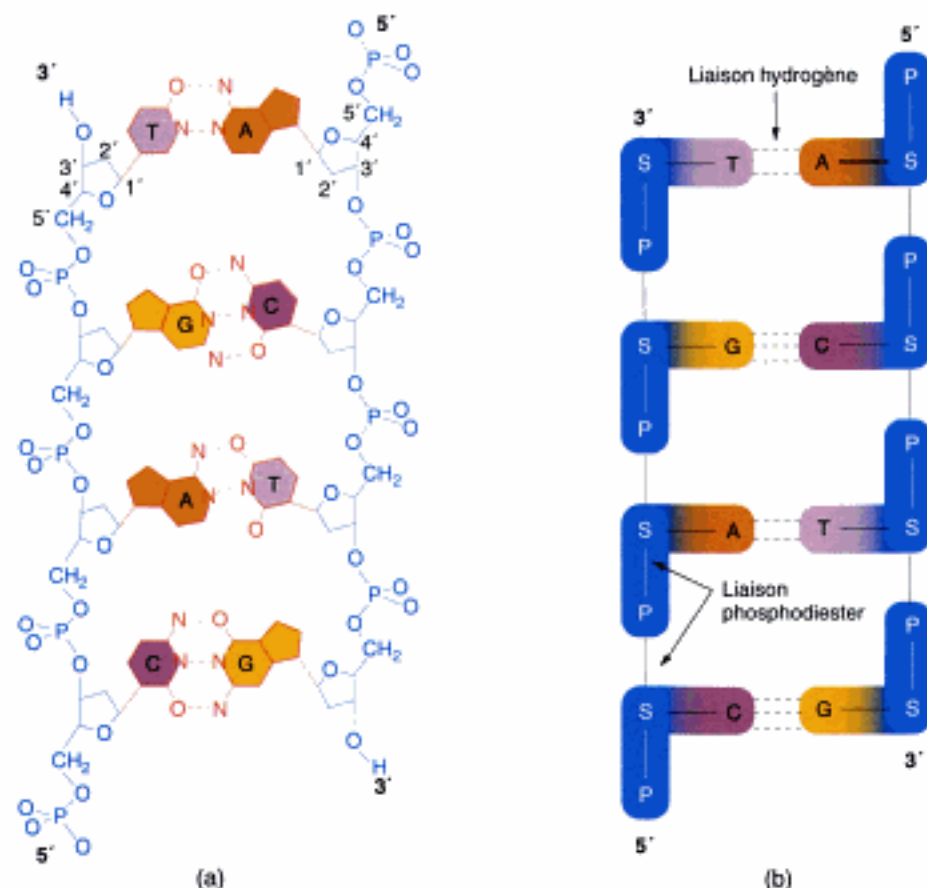
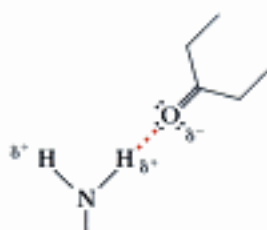


Figure 2-2 L'organisation des composants de l'ADN. Un segment de la double hélice a été déroulé pour présenter plus clairement les structures. (a) Un schéma précis des constituants chimiques montrant le squelette sucre-phosphate en bleu et les liaisons hydrogène entre les bases au centre de la molécule. (b) Une version simplifiée du même segment, soulignant l'arrangement antiparallèle des nucléotides, qui sont représentés sous la forme de structures en L avec des « orteils » de phosphate en 5' et des « talons » de sucre en 3'.

les deux brins sans déformer fondamentalement la molécule d'ADN double-brin sont A•T et G•C.

L'association de A avec T et de G avec C se fait par l'intermédiaire de **liaisons hydrogène**. Le dessin ci-dessous est un exemple de liaison hydrogène :



Chaque atome d'hydrogène du groupement NH_2 est légèrement positif (δ^+) car l'atome d'azote a tendance à attirer les électrons impliqués dans la liaison N-H, laissant ainsi l'atome d'hydrogène légèrement dépourvu d'électrons. L'atome d'oxygène possède six électrons libres dans sa couche externe, ce qui le rend légèrement négatif (δ^-). Une liaison hydrogène se forme entre un H légèrement positif et un atome légèrement négatif, dans cet exemple, O. Les liaisons hydrogène sont très faibles (seulement 3 pour cent de la force d'une liaison covalente), mais cette faiblesse (comme nous le verrons) est importante pour le rôle de l'ADN dans l'hérédité. Une autre caractéristique chimique est importante : la liaison hydrogène est bien plus forte si les atomes participants « se regardent » (c'est-à-dire, si leurs liaisons sont alignées), comme on le voit dans le schéma.

Il est à noter que puisque la paire G•C possède trois liaisons hydrogène, alors que la paire A•T n'en contient que deux, on pourrait penser qu'un ADN contenant de nombreuses paires G•C serait plus stable qu'un ADN contenant de nombreuses paires A•T. Cette hypothèse est en effet confirmée. La chaleur entraîne la séparation de la double hélice d'ADN (un processus appelé *fusion de l'ADN* ou *dénaturation de l'ADN*). On peut montrer que les ADN ayant un contenu plus élevé en G+C requièrent des températures plus élevées pour être dénaturés.

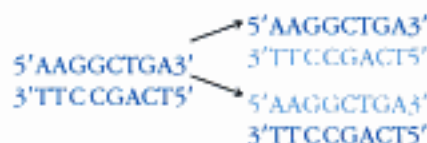
Bien que les liaisons hydrogène considérées individuellement soient faibles, les deux brins de la molécule d'ADN sont maintenus ensemble de façon relativement stable car ces liaisons sont présentes en grand nombre. Il est important que les brins soient associés par l'intermédiaire de ces interactions faibles, car ils doivent se séparer au cours de la réplication de l'ADN et au cours de sa transcription en ARN.

Les deux brins nucléotidiques appariés adoptent automatiquement une configuration en double hélice (figure 2-3), essentiellement à cause de l'interaction des paires de bases. Les paires de bases, qui sont des structures planes, s'empilent les unes sur les autres au centre de la double hélice. L'empilement (figure 2-3c) ajoute à la stabilité de la molécule d'ADN en excluant les molécules d'eau des espaces compris entre les paires de bases. La

forme la plus stable qui résulte de l'empilement des bases est une double hélice avec des sillons de deux tailles différentes qui s'enroulent en une spirale. Il s'agit du grand sillon et du petit sillon, que l'on peut voir dans les modèles. Un brin unique de nucléotides n'a pas de structure hélicoïdale ; cette forme en hélice de l'ADN dépend entièrement de l'appariement et de l'empilement des bases dans les brins antiparallèles.

La structure de l'ADN reflète sa fonction

Comment la structure de l'ADN remplit-elle les exigences d'une molécule héréditaire ? En premier lieu, la duplication. Avec l'orientation antiparallèle des brins d'ADN et les règles régissant un appariement correct, nous pouvons imaginer de quelle façon un ADN peut être correctement dupliqué : chaque brin sert de **matrice** (guide d'alignement) pour la synthèse de son brin complémentaire. Si par exemple, un brin a la séquence AAGGCTGA (lue dans le sens 5' vers 3'), alors on sait automatiquement que son brin complémentaire aura la séquence TTCCGACT (dans le sens 3' vers 5'). La réplication est basée sur cette règle simple. Les deux brins d'ADN se séparent et chacun sert de matrice pour la construction d'un nouveau brin complémentaire.



Une enzyme appelée *ADN polymérase* est responsable de la construction des nouveaux brins d'ADN, sélectionnant chaque base du nouveau brin pour qu'elle soit complémentaire de la base correspondante dans l'ancien brin, le brin matrice. La complémentarité des brins d'ADN sous-tend donc la totalité du processus de duplication exacte. Nous décrirons plus précisément ce processus au cours du chapitre 4.

La deuxième exigence concernant l'ADN est d'avoir un contenu informationnel. Ce besoin informationnel est rempli par sa séquence nucléotidique, qui ressemble à un langage écrit. La troisième exigence, la mutation, est simplement le remplacement occasionnel, la délétion ou l'addition d'une ou plusieurs paires de nucléotides, entraînant un changement de l'information codée.

MESSAGE

L'ADN double-brin est composé de deux chaînes nucléotidiques entrecroisées et antiparallèles, constituées chacune d'un squelette sucre-phosphate dont les bases sont liées par des liaisons hydrogène aux bases complémentaires de l'autre chaîne.

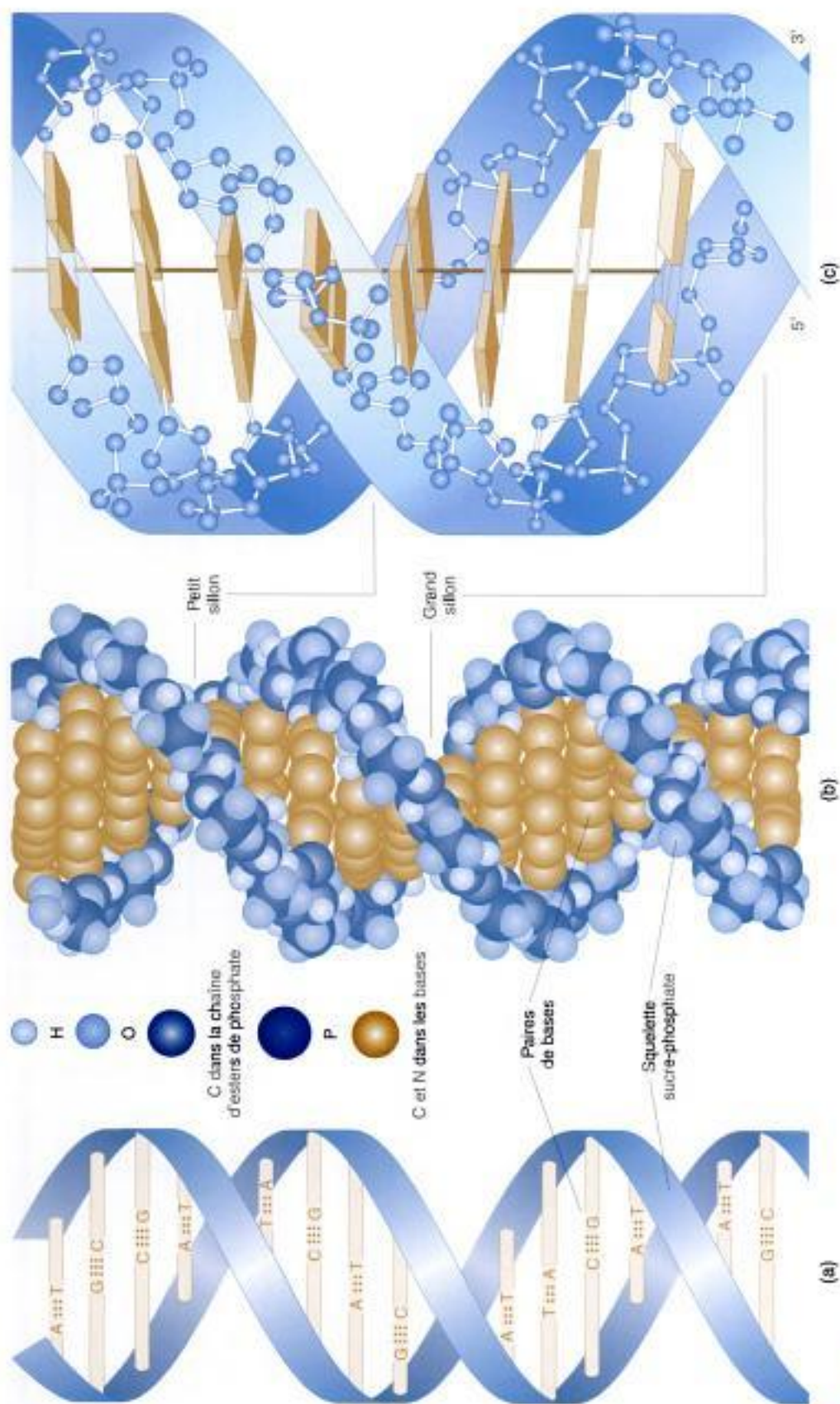


Figure 2-3 Trois représentations de la double hélice d'ADN.

LA NATURE DES GÈNES

La taille et la fonction des gènes sont diverses. Cependant, on peut définir pour la plupart des gènes certaines caractéristiques topographiques.

Les principales régions d'un gène

Au chapitre 1, un gène a été défini simplement comme une région d'ADN capable d'être transcrite en une molécule d'ARN fonctionnel. Il faut ajouter à cela que le gène doit non seulement pouvoir être transcrit en un ARN fonctionnel (pour la plupart des gènes, un ARNm) mais que cet ARN doit également être fabriqué au bon endroit et au bon moment lors du développement de l'organisme. C'est seulement dans ce cas qu'un gène est parfaitement fonctionnel. Pour que cela soit possible, à l'une des extrémités du gène se trouve une **région régulatrice**, un segment d'ADN constitué d'une séquence nucléotidique spécifique qui lui permet de recevoir et de répondre à des signaux provenant d'autres parties du génome ou de l'environnement. En dernier lieu, ces signaux d'activation sont convertis en protéines régulatrices qui se fixent à la région régulatrice du gène et initient la transcription dans la région adjacente, qui code l'ARN. À l'autre extrémité du

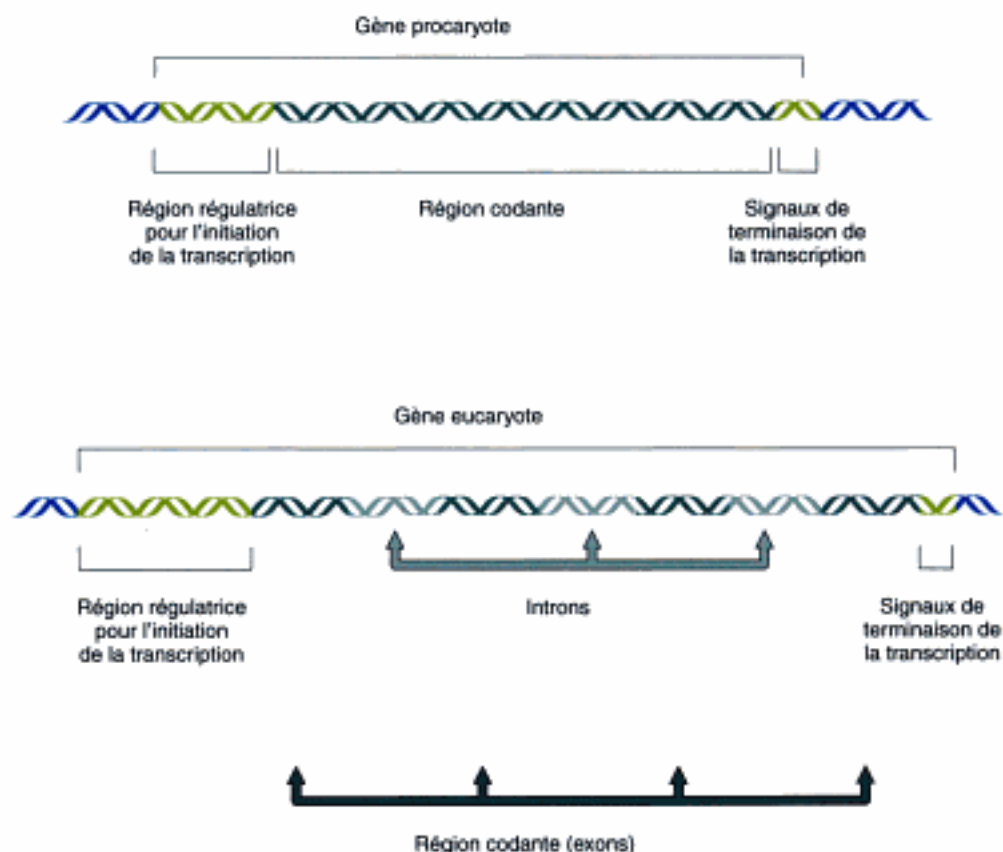
gène se trouve une région qui contient les signaux pour terminer le transcrit.

MESSAGE

Un gène est une région d'ADN chromosomique qui peut être transcrite en un ARN fonctionnel au bon endroit et au bon moment au cours du développement.

Essayer de préciser davantage ce que sont les gènes est compliqué par le fait que (comme nous l'avons vu au chapitre 1), de nombreux gènes eucaryotes contiennent de mystérieux segments d'ADN, appelés **introns**, intercalés dans la région transcrite du gène. Les introns ne contiennent aucune information de séquence concernant les produits fonctionnels des gènes, tels que les protéines. Ils sont transcrits en même temps que les régions codantes (appelées **exons**), mais ils sont ensuite excisés du transcrit initial. Puisqu'une séquence correcte dans les introns (ainsi que dans la région régulatrice) est nécessaire pour produire un transcrit de la taille adéquate, au bon moment et au bon endroit, les introns (comme les régions codantes et régulatrices) doivent être considérés comme des parties de l'unité fonctionnelle globale — en d'autres termes, comme faisant partie du gène (figure 2-4).

Figure 2-4 La structure générale d'un gène chez les Procaryotes et les Eucaryotes. La région codante (vert foncé) contient l'information concernant la structure du produit du gène (généralement une protéine). Les régions régulatrices adjacentes (vert citron) contiennent des séquences reconnues par des protéines qui s'y fixent alors, soit pour fabriquer l'ARN correspondant au gène, soit pour influencer la quantité d'ARN fabriqué. Remarquez que dans les gènes eucaryotes, la région codante est souvent interrompue par un ou plusieurs introns non codants (vert pâle).



La figure 2-5 présente le nombre d'exons (égal au nombre d'introns plus un) dans des échantillons de gènes provenant de trois organismes eucaryotes, la levure de bière (*Saccharomyces cerevisiae*), la mouche du vinaigre (*Drosophila melanogaster*) et les mammifères. Comme vous pouvez le constater, il existe des différences frappantes. Chez les mammifères, les gènes sont généralement très longs et les unités de transcription comportent de nombreux introns. Le tableau 2-1 montre les tailles moyennes des gènes codant des protéines chez des organismes représentatifs et les compare à la taille de l'ARN messager (qui a la même taille que tous les exons combinés). Les régions codantes des gènes de tous ces organismes sont comprises entre 2 et 3 kb, mais les tailles des gènes sont très différentes — une grande partie de cette différence est liée au nombre moyen d'introns.

Les gènes sont entourés d'une autre sorte d'ADN

Le séquençage nucléotidique et d'autres études physiques ont montré qu'entre les gènes se trouve de l'ADN, dont on ignore la fonction la plupart du temps. La taille et la nature de cet ADN varient selon le génome. Par exemple chez les champignons, il est présent en faible quantité, alors que chez les mammifères, les régions intergéniques peuvent être longues. Certains caractères topographiques généraux des chromosomes sont illustrés dans la figure 2-6. Quelque part entre ces gènes et cet ADN extragénique se trouvent des séquences d'ADN, situées parfois très loin d'un gène donné, qui peuvent affecter la régulation de ce gène. On pourrait donc les considérer comme des parties de l'unité fonctionnelle du gène, même si elles en sont

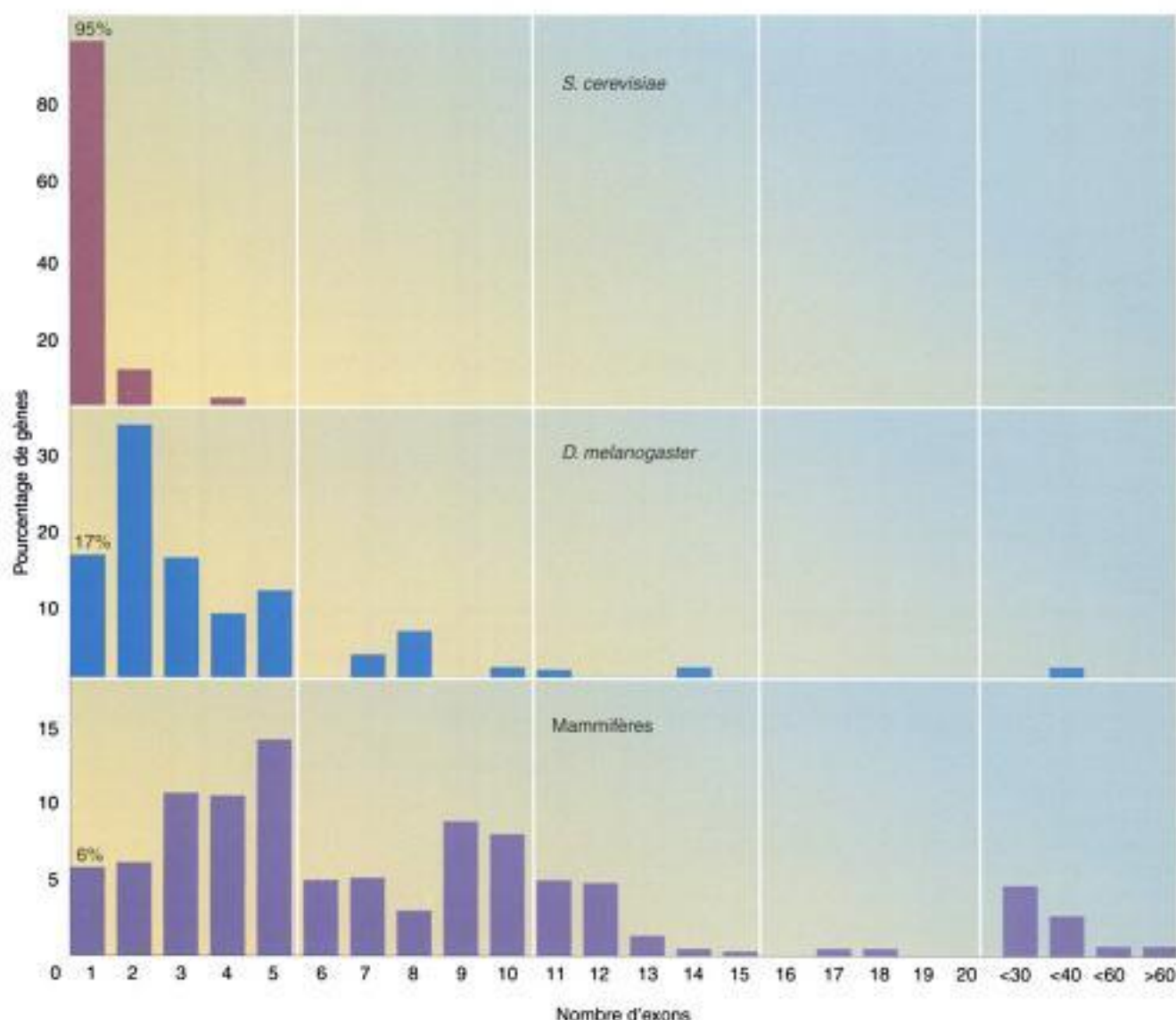


Figure 2-5 Distribution du nombre d'exons dans les gènes de trois organismes.

TABLEAU 2-1
La relation entre la taille des gènes et la taille des ARNm

Espèce	Nombre moyen d'exons	Longueur moyenne des gènes (en kb)	Longueur moyenne des ARNm (en kb)
<i>Hemophilus influenzae</i>	1	1,0	1,0
<i>Methanococcus jannaschii</i>	1	1,0	1,0
<i>S. cerevisiae</i>	1	1,6	1,6
Champignon filamenteux	3	1,5	1,5
<i>C. elegans</i>	4	4,0	3,0
<i>D. melanogaster</i>	4	11,3	2,7
Poulet	9	13,9	2,4
Mammifères	7	16,6	2,2

Source : D'après B. Lewin, Gènes VI, Tableau 21-1. Éditions DeBoeck Université

séparées par de longs segments d'ADN qui n'ont rien à voir avec le gène en question. Chez de nombreux Eucaryotes, une partie de l'ADN situé entre les gènes est répétitif. Il est constitué de plusieurs types différents d'unités répétées dans l'ensemble du génome. Une partie de cet ADN répétitif est dispersé ; une autre se rencontre sous la forme de séquences « en tandem » contiguës. On trouve également de l'ADN répétitif dans certains introns. La proportion de cet ADN est distincte pour des espèces différentes et de plus, il y a une variation du nombre de répétitions à l'intérieur d'une même espèce. La fonction de l'ADN répétitif est toujours un mystère, mais il existe plusieurs hypothèses quant à son origine. (Ces sujets seront traités en détail au chapitre 12.)

LA NATURE DES GÉNOMES

Tous les organismes ont un génome comportant de l'ADN et tous possèdent des gènes. Mais derrière ces généralités apparaissent de nombreuses questions sur ces génomes. Quelle taille ont-ils ? Combien de gènes contiennent-ils et comment ce nombre varie-t-il d'une espèce à l'autre ? Combien de chromosomes les génomes contiennent-ils et de quelle façon les gènes sont-ils regroupés sur les chromosomes ?

Comprendre l'organisation d'un génome est également important pour comprendre la transmission des gènes. Si l'on considère la transmission des gènes lors de

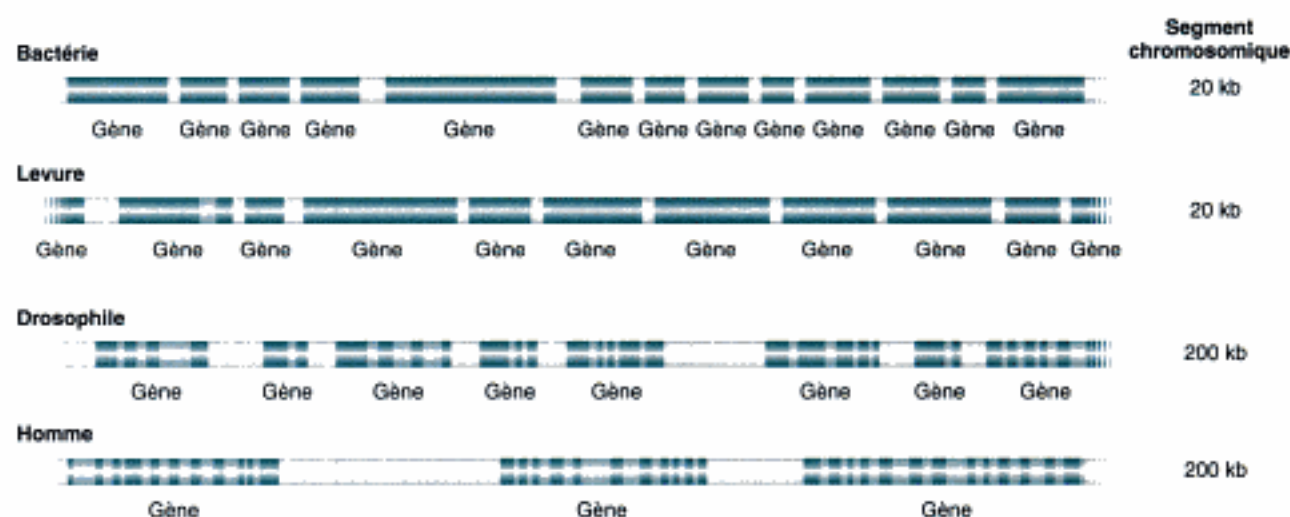


Figure 2-6 Topographie des gènes chez quatre organismes. Vert clair = introns ; vert foncé = exons ; blanc = régions intergéniques.

la division cellulaire, ceux-ci doivent être dupliqués et les copies dupliquées doivent ségréger vers les deux cellules filles. Le mécanisme qui gouverne la transmission des gènes sera une des questions essentielles de ce livre. Pour le moment, il est simplement important de réaliser que de nombreux principes de la transmission génétique viennent du fait que les gènes sont portés par les chromosomes, et découlent de la structure des chromosomes eux-mêmes, car ce sont ces structures qui doivent être répliquées et réparties entre les deux cellules filles. Au niveau moléculaire, les chromosomes sont des structures complexes contenant à la fois de l'ADN et des protéines, mais du point de vue du généticien, le point essentiel est qu'un chromosome contient une seule molécule d'ADN double-brin.

La taille du génome

L'une des questions essentielles qui demande une réponse est « De quelle taille sont les génomes ? ». Les tailles des génomes sont mesurées en unités de milliers de paires de nucléotides (appelées *kilobases*, kb) ou en millions de paires de nucléotides (*mégabases*, Mb). La figure 2.7 présente la gamme des tailles de génomes représentatifs. Remarquez que la taille des génomes augmente généralement avec la complexité du groupe, mais qu'il y a des variations considérables (jusqu'à un millier de fois !) dans certains groupes. Le tableau 2-2 compare la taille des génomes représentatifs et des composants génomiques. Le tableau indique également le nombre de gènes pour ces exemples. Le nombre de gènes est grossièrement proportionnel à la taille du génome, mais il y a des écarts dus à

l'ADN répétitif et aux introns. Ce sont les virus et les plasmides, tous deux incapables d'une existence indépendante, qui comportent le plus petit nombre de gènes.

Nous allons maintenant examiner en détail les différents génomes dans l'ordre croissant de leur taille.

Les génomes plasmidiques

Les cellules bactériennes isolées dans la nature contiennent souvent de petits éléments d'ADN qui ne sont pas essentiels au fonctionnement élémentaire de la cellule. On appelle ces éléments, des *plasmides*. Les plasmides sont des molécules symbiotiques qui ne peuvent survivre hors des cellules. Même si les plasmides n'appartiennent pas au système opérationnel de base de leur cellule hôte, certains sont très complexes et portent de nombreux gènes. Il est donc approprié de qualifier leur ADN de « génome plasmidique ». Les plasmides bactériens contiennent souvent des gènes extrêmement utiles à l'hôte bactérien. Certains initient par exemple la fusion des cellules bactériennes, confèrent une résistance à un antibiotique ou encore produisent des toxines.

On rencontre aussi occasionnellement des plasmides dans des cellules de plantes ou de champignons. On trouve la plupart des plasmides à l'intérieur des mitochondries et des chloroplastes, mais il y en a parfois dans le noyau ou le cytosol. Au contraire des plasmides bactériens mentionnés plus haut, ces plasmides eucaryotes n'apportent apparemment aucun avantage à leur hôte — ils semblent vivre égoïstement, et se consacrer entièrement à leur propagation.

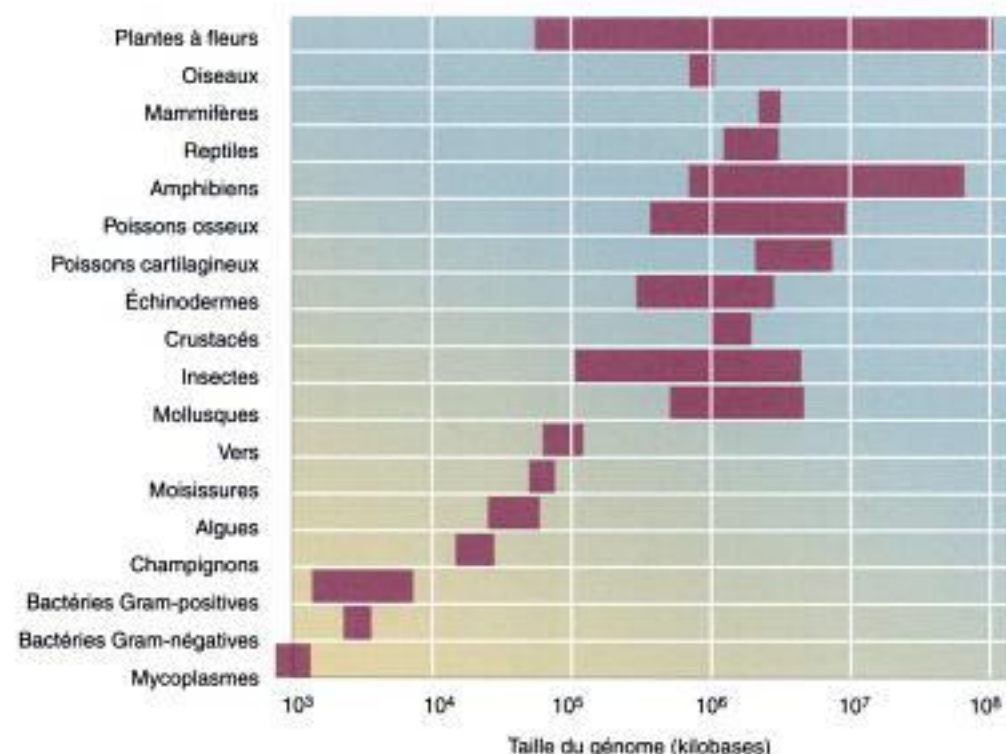


Figure 2-7 Quantité d'ADN par génome chez différents organismes.

TABLEAU 2-2
Génomes : taille et nombre de gènes

Génome	Groupe	Taille (kb)	Nombre de gènes
<i>Noyau eucaryote</i>			
<i>Saccharomyces cerevisiae</i>	Levure	13 500 (L)	6 000
<i>Caenorhabditis elegans</i>	Nématode	100 000 (L)	13 500
<i>Arabidopsis thaliana</i>	Végétal	120 000 (L)	25 000
<i>Homo sapiens</i>	Humain	3 000 000 (L)	100 000
<i>Procaryote</i>			
<i>Escherichia coli</i>	Bactérie	4 700 (C)	4 000
<i>Hemophilus influenzae</i>	Bactérie	1 830 (C)	1 703
<i>Methanococcus jannaschii</i>	Bactérie	1 660 (C)	1 738
<i>Virus</i>			
T4	Virus bactérien	172 (L/C)	300
HCMV (groupe des herpès)	Virus humain	229 (L)	200
<i>Organites eucaryotes</i>			
Mitochondrie de <i>S. cerevisiae</i>	Levure	78 (C)	34
Mitochondrie de <i>H. sapiens</i>	Humain	17 (C)	37
Chloroplaste de <i>Marchantia polymorpha</i>	Bryophyte de la classe des Hépatiques	121 (C)	136
<i>Plasmides</i>			
Plasmide F	Chez <i>E. coli</i>	100 (C)	29
kalilo	Chez <i>Neurospora</i> , un champignon	9 (L)	2

Remarque : C = circulaire ; L = linéaire ; L/C = linéaire dans les virus libres, circulaire dans les cellules.

Pour leur réplication et leur entretien, les plasmides dépendent de la machinerie cellulaire codée par le génome hôte. Les plasmides bactériens sont le plus souvent circulaires, mais il en existe aussi des types linéaires. Chez les champignons et les plantes, les plasmides linéaires sont les plus courants, mais on connaît aussi des types circulaires chez les champignons.

L'ADN des organites

La majeure partie du génome eucaryote est contenue dans les chromosomes du noyau. Cependant, en plus de l'ADN nucléaire, certains organites cellulaires — mitochondries et chloroplastes — contiennent également un type de chromosome qui leur est spécifique. Les chromosomes mitochondriaux et chloroplastiques sont des molécules d'ADN double-brin. Les mitochondries et les chloroplastes contiennent de multiples copies de leurs chromosomes, et dans chaque cellule, ces organites sont présents en nombre variable. Le nombre de copies de ces chromosomes par cellule peut donc être très élevé, souvent de l'ordre de centaines. Il est donc relativement facile d'extraire l'ADN des fractions contenant les organites de cellules broyées. Les chromosomes des organites contiennent des gènes spécifiques des fonctions de l'orga-

nite concerné. Malgré cela, la plupart des fonctions biologiques qui se déroulent à l'intérieur de ces organites sont spécifiées par des gènes du génome nucléaire. Les génomes des organites et le génome nucléaire n'ont aucun élément commun. Les mitochondries et les chloroplastes étaient probablement à l'origine des cellules procaryotes qui pénétrèrent dans une autre cellule et établirent une relation de symbiose avec celle-ci (comme le propose la théorie endosymbiotique de l'origine des Eucaryotes). Au cours de l'évolution, la plupart des gènes procaryotes originaux furent transférés au génome nucléaire ou perdus. Les génomes mitochondriaux peuvent être éliminés de certains organismes tels que les levures (qui peuvent produire leur énergie grâce aux réactions de fermentation), mais la plupart des organismes ne peuvent survivre sans eux. Il y a donc toujours une interdépendance entre les parties du génome situées dans les organites et celles présentes dans le noyau. Les chloroplastes peuvent être éliminés seulement chez les organismes photosynthétiques qui peuvent survivre en captant les nutriments préformés de leur environnement (c'est-à-dire, ceux qui peuvent se comporter comme des hétérotrophes). L'analyse des génomes d'organites montre qu'ils sont normalement circulaires, mais on peut détecter des formes linéarisées de ces génomes dans certaines cellules. Un grand nombre des gènes d'organites contiennent des introns.

Les génomes viraux

Un virus est une particule non vivante qui ne peut se reproduire qu'en infectant une cellule vivante et en détournant la machinerie cellulaire de la cellule hôte pour engendrer des particules virales filles. Les virus des bactéries sont également appelés *bactériophages* (littéralement, « mangeurs de bactéries »). Les virus sont constitués d'une enveloppe protéique et d'une partie centrale qui contient leur génome. Au cours de l'infection, le génome viral pénètre dans la cellule, soit grâce à une fusion de la paroi virale avec la membrane plasmique de la cellule, soit par un processus qui ressemble à une injection par une seringue.

Il y a donc une phase intracellulaire dans le cycle de réplication du génome viral, ainsi qu'une phase au cours de laquelle ce génome est empaqueté en particules virales infectieuses. Comme il est plus facile de purifier le génome viral lorsqu'il se trouve à l'intérieur de la particule virale extracellulaire, les génomes viraux sont généralement décrits d'après les structures adoptées dans cette particule virale. Les génomes viraux sont très variés. Les génomes de la plupart des virus sont constitués d'ADN. Lorsqu'il est empaqueté dans la particule virale, l'ADN de certains virus est double-brin, dans d'autres cas, simple-brin. D'autres virus pourtant, tels que les rétrovirus (dont le VIH est un exemple) ont des génomes d'ARN dans la particule virale. Là encore, ces génomes d'ARN peuvent être double-brin ou simple-brin. Certains génomes viraux contiennent des molécules d'ADN ou d'ARN linéaire, tandis que d'autres sont circulaires.

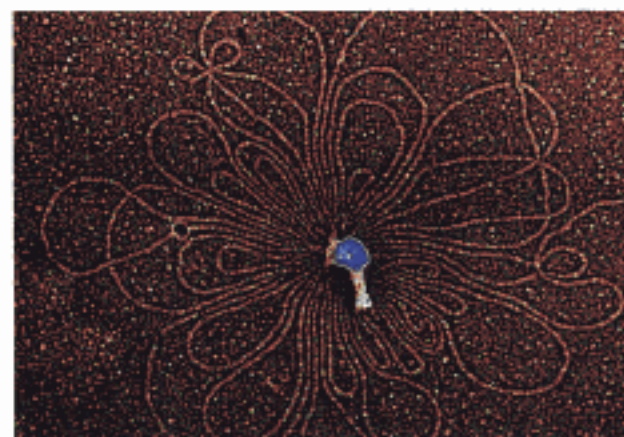
On observe donc une diversité bien plus grande des types de molécules qui forment les génomes viraux, par rapport aux génomes d'ADN double-brin que l'on rencontre dans toutes les cellules vivantes. Cette diversité est basée en partie sur les différentes stratégies qui permettent d'empaquer le génome dans les particules virales infectieuses et en partie sur la diversité de leurs histoires au cours de l'évolution. Sans tenir compte de la nature du génome viral empaqueté dans la particule infectieuse, il y a une phase du cycle viral dans la cellule pendant laquelle le génome viral est converti en une molécule d'ADN double-brin, c'est-à-dire de même nature moléculaire que le génome de l'hôte. Selon le virus et les conditions qui règnent dans la cellule, cette forme intracellulaire du génome peut être incorporée dans un chromosome de l'hôte ou bien exister sous la forme d'une molécule séparée. La plupart des gènes viraux ne contiennent pas d'introns.

Dans les petits génomes des plasmides, des organites et des virus, les gènes sont proches : les espaces intergéniques sont relativement courts. C'est également vrai pour les génomes procaryotes mais cela contraste avec les chromosomes de nombreux animaux et végétaux, dans lesquels (comme nous le verrons) il y a souvent de longues étendues d'ADN entre les gènes.

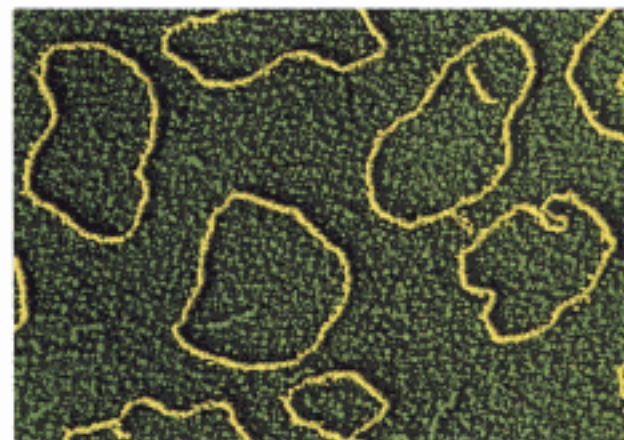
Au contraire des génomes des organismes, les petits génomes n'ont pas de grandes quantités de protéines associées. Les virus linéaires et les plasmides linéaires ont

une protéine spéciale, attachée en permanence au groupement 5' phosphate de chaque extrémité. Ces protéines jouent un rôle essentiel lorsque l'ADN se réplique.

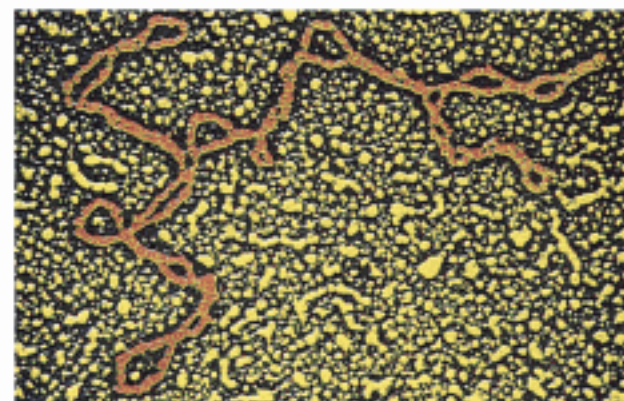
Des micrographies électroniques de petits génomes représentatifs sont présentées dans la figure 2-8.



(a)



(b)



(c)

Figure 2-8 Micrographies électroniques de quelques petits génomes. (a) phage T2 ; (b) plasmides d'*E. coli* ; (c) ADN mitochondrial humain. (a) Biology Media/Science Photo Library/Photo Researchers. (b) Dr Gopal Murti/Science Photo Library/Photo Researchers. (c) NRL/Science Photo Library/Photo Researchers.

Les génomes procaryotes

Le génome de la plupart des Procaryotes est contenu dans un seul chromosome. Dans presque tous les cas, ce chromosome est une double hélice d'ADN circulaire fermée. Il existe quelques exceptions, telles que la bactérie *Borrelia burgdorferi*, chez laquelle le chromosome est une double hélice unique d'ADN linéaire. Certains génomes bactériens sont constitués de plusieurs chromosomes différents.

Les gènes bactériens sont assez proches les uns des autres, avec assez peu d'espace intergénique et les introns sont extrêmement rares. Dans certaines régions des génomes procaryotes, certains gènes ayant une parenté fonctionnelle sont regroupés et une seule molécule d'ARN est fabriquée à partir de l'unité entière ; on appelle cette unité un *opéron*. Les gènes intervenant dans l'utilisation du lactose chez la bactérie du côlon *Escherichia coli* constituent un opéron et sont transcrits sous la forme d'une seule molécule d'ARNm. Les opérons sont très rares chez les Eucaryotes. Les propriétés fonctionnelles des opérons seront examinées au chapitre 14. Dans chaque cellule bactérienne il peut y avoir une ou plusieurs copies identiques de l'unique chromosome.

Dans les micrographies électroniques des cellules bactériennes, on voit l'ADN arrangé en une masse dense appelée un **nucléotide**. Lorsque les cellules sont cassées,

l'emballage de ce nucléotide est perdu et l'ADN s'échappe en un écheveau désorganisé (figure 2-9). Les génomes bactériens sont associés à des protéines qui semblent aider à l'emballage du génome en un nucléotide, mais les fonctions précises de ces protéines associées à l'ADN ne sont pas comprises.

MESSAGE

Une cellule procaryote possède une ou plusieurs copies d'un génome circulaire unique et parfois, une ou plusieurs copies de plasmides.

Les génomes nucléaires eucaryotes

Chez les organismes eucaryotes, on trouve la grande majorité des gènes dans les chromosomes du noyau (figure 2-10). La plupart des espèces eucaryotes sont classifiées en **diploïdes**, portant deux jeux de chromosomes nucléaires (deux copies de la fraction nucléaire du génome) dans chaque noyau des cellules du corps, ou **haploïdes**, avec un seul jeu de chromosomes par noyau. La majorité des champignons et des algues sont haploïdes, tandis que la plupart des autres Eucaryotes, y compris les animaux et les plantes à fleurs, sont diploïdes. Cependant, il faut remarquer que les organismes diploïdes produisent des cellules reproductrices haploïdes (telles que les ovules et les spermatozoïdes chez les animaux) et qu'inversement, les organismes haploïdes produisent, au cours de la phase sexuelle de leur cycle vital, des cellules diploïdes spécialisées.

La lettre **n** est utilisée pour désigner le nombre de chromosomes dans un génome nucléaire. Un état haploïde est donc désigné par *n* (c'est-à-dire $1 \times n$) et un état



Figure 2-9 Micrographie électronique du génome de la bactérie *E. coli*, libéré de la cellule par un choc osmotique. (Dr Gopal Murti/Science Photo Library/Photo Researchers.)

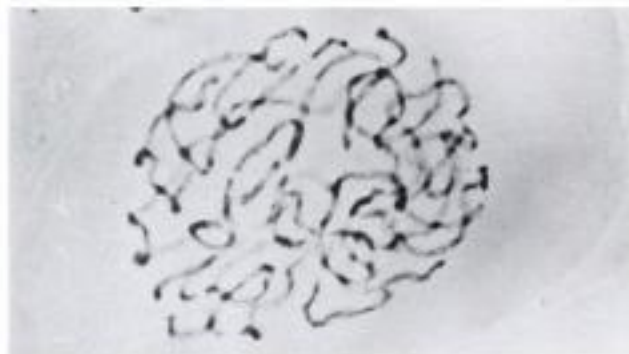


Figure 2-10 Les chromosomes dans un noyau eucaryote. Les chromosomes sont sous une forme étirée dans les états actifs du cycle cellulaire mais se contractent en vue de la division cellulaire. Cette photographie qui montre des chromosomes commençant à se compacter permet néanmoins de se représenter leur état étiré. (Anthony Griffiths.)

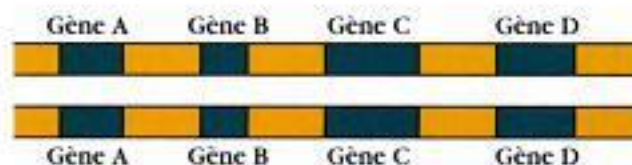
diploïde par $2n$ (c'est-à-dire $2 \times n$). Le symbole n est appelé le **nombre haploïde de chromosomes**. On connaît également des états $3n$, $4n$, $5n$, $6n$, etc. en particulier chez les végétaux ; on les qualifie de polyploïdes. Nous y reviendrons au chapitre 8. Le nombre de jeux de chromosomes (1, 2, 3, 4, 5, 6, etc.) est parfois appelé la **ploïdie** ou le **niveau de ploïdie**. Notez que par convention le niveau de ploïdie fait référence au nombre de jeux de chromosomes dans une cellule qui ne s'est pas engagée dans la division cellulaire.

MESSAGE

Le nombre de chromosomes dans un noyau eucaryote est déterminé par le nombre de chromosomes dans le jeu de chromosomes (le nombre haploïde de chromosomes), multiplié par le nombre de jeux (niveau de ploïdie).

Dans une cellule diploïde, puisqu'il y a deux jeux de chromosomes, il existe deux chromosomes de chaque type — deux exemplaires du chromosome 1, deux du chromosome 2, deux du chromosome 3 et ainsi de suite. Les membres d'une paire sont appelés **chromosomes homologues** ou simplement **homologues**. Les membres d'une paire d'homologues sont globalement identiques en ce qui concerne leur taille et les gènes qu'ils contiennent. Ils portent les mêmes gènes dans des positions relatives

identiques. On peut donc représenter comme ceci un segment d'une paire de chromosomes homologues dans une cellule diploïde



Toutefois, une paire d'homologues peut porter des allèles différents d'un même gène (voir la définition d'un *allèle* dans le chapitre 1), ce qui entraînera des différences mineures entre les homologues. Le nombre de chromosomes est déterminé simplement en réalisant des comptages sur des préparations colorées, observées sous un microscope photonique. Les chromosomes sont plus faciles à voir lorsqu'ils sont dans une forme condensée, comme pendant la division cellulaire. Le nombre de chromosomes de quelques organismes diploïdes est indiqué dans le tableau 2-3.

Des chromosomes distincts dans un génome contiennent des gènes différents. On peut démontrer cela en examinant des individus anormaux qui possèdent un chromosome supplémentaire. Le chromosome supplémentaire

TABEAU 2-3
Nombre de paires de chromosomes chez
différentes espèces de végétaux et d'animaux

Nom usuel	Espèces	Nombre de paires de chromosome	Nom usuel	Espèces	Nombre de paires de chromosome
Moustique	<i>Culex pipiens</i>	3	Blé	<i>Triticum aestivum</i>	21
Mouche commune	<i>Musca domestica</i>	6	Homme	<i>Homo sapiens</i>	23
Oignon	<i>Allium cepa</i>	8	Pomme de terre	<i>Solanum tuberosum</i>	24
Crapaud	<i>Bufo americanus</i>	11	Bovin	<i>Bos taurus</i>	30
Riz	<i>Oryza sativa</i>	12	Âne	<i>Equus asinus</i>	31
Grenouille	<i>Rana pipiens</i>	13	Cheval	<i>Equus caballus</i>	32
Alligator	<i>Alligator mississippiensis</i>	16	Chien	<i>Canis familiaris</i>	39
Chat	<i>Felis domesticus</i>	19	Poulet	<i>Gallus domesticus</i>	39
Souris commune	<i>Mus musculus</i>	20	Carpe	<i>Cyprinus carpio</i>	52
Singe rhésus	<i>Macaca mulatta</i>	21			

induit une apparence anormale, caractéristique de ce chromosome, ce qui suggère que chaque chromosome est différent. La figure 2-11 présente un exemple de ces observations chez des plantes. La cartographie (chapitre 5) et le séquençage (chapitre 12) des chromosomes ont confirmé l'idée que des chromosomes différents portent des gènes distincts. Comme nous le verrons dans le paragraphe suivant, l'aspect des chromosomes est très variable lorsqu'on les observe sous le microscope. Ces caractéristiques propres à chaque chromosome suggèrent également fortement que le contenu génique des différents chromosomes varie.

MESSAGE

Une cellule eucaryote possède un ou deux jeux de chromosomes nucléaires linéaires, de multiples copies d'un chromosome mitochondrial circulaire, de multiples copies d'un chromosome chloroplastique circulaire (chez les végétaux seulement) et parfois de multiples copies de plasmides (chez certains champignons et plantes).

La figure 2-12 à la page suivante est un résumé schématique des principales caractéristiques de taille et d'organisation génétique des différents types de génomes dont nous avons parlé jusqu'ici.

LA NATURE DES CHROMOSOMES NUCLÉAIRES EUCARYOTES

L'ensemble des chromosomes tel qu'on le voit sous le microscope est appelé un **caryotype** (« type nucléaire »). Le caryotype est défini par le nombre de chromosomes et par d'autres caractéristiques visibles.

Les caractéristiques chromosomiques visibles

L'étude microscopique des chromosomes et l'analyse de leurs propriétés génétiques est appelée *cytogénétique*, une discipline qui combine cytologie et génétique. Puisqu'une grande partie de ce livre sera consacrée aux propriétés des chromosomes, nous commencerons par un aperçu de la morphologie visible des chromosomes, des caractéristiques que l'on utilise couramment en analyse cytogénétique pour trier les chromosomes.

On peut voir les chromosomes relativement facilement sous le microscope, mais seulement juste avant, pendant et juste après la division cellulaire. Lorsqu'une cellule se divise, le noyau et ses chromosomes se divisent également. La division cellulaire et nucléaire fera l'objet du chapitre 4. Pour l'instant, il nous faut remarquer seulement que les chromosomes doivent devenir très compacts au cours de la division cellulaire, car ils subissent de nombreux mouvements au cours de leur répartition entre les cellules filles. Donc, les composants d'un chromosome (y compris son ADN) passent de la forme étirée que l'on observe dans les cellules qui ne sont pas en cours de division (c'est-à-dire qui sont en interphase) à une forme plus courte, plus épaisse qui peut facilement être manipulée par l'appareil de division de la cellule. Ce sont les caractéristiques de ces chromosomes *condensés* que l'on observe sous le microscope, mais on peut supposer que ces caractéristiques ont les mêmes positions relatives sur les chromosomes *étirés* des cellules en interphase.

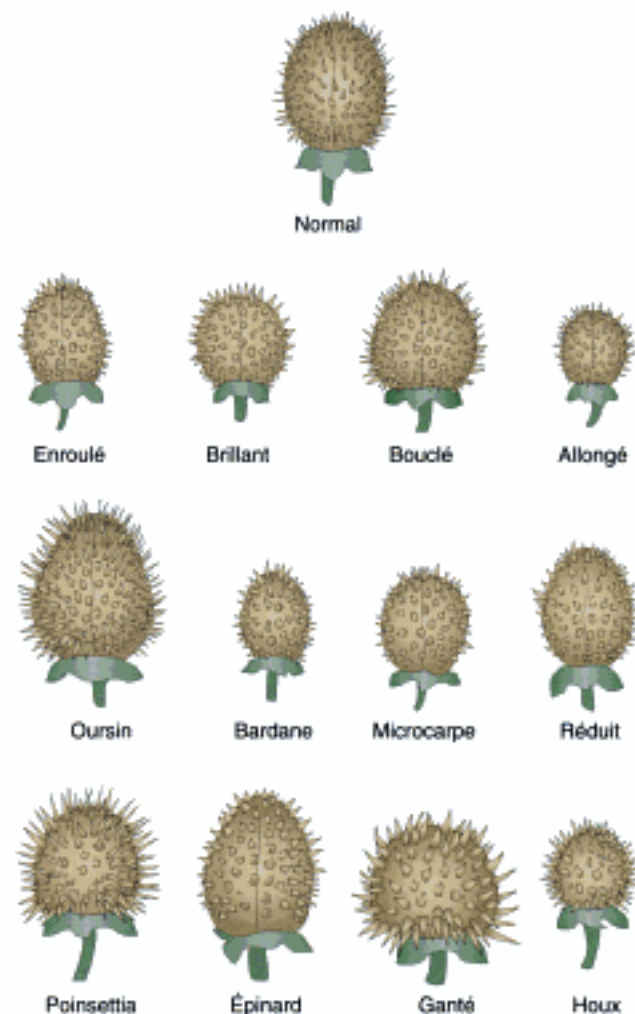


Figure 2-11 Des fruits de la plante *Datura*, qui possèdent tous un chromosome supplémentaire distinct. Leur apparence caractéristique suggère que chaque chromosome est différent. (D'après E. W. Sinnott, L. C. Dunn et T. Dobzhansky, *Principles of Genetics*, 5e édition. McGraw-Hill Book Company, New York.)

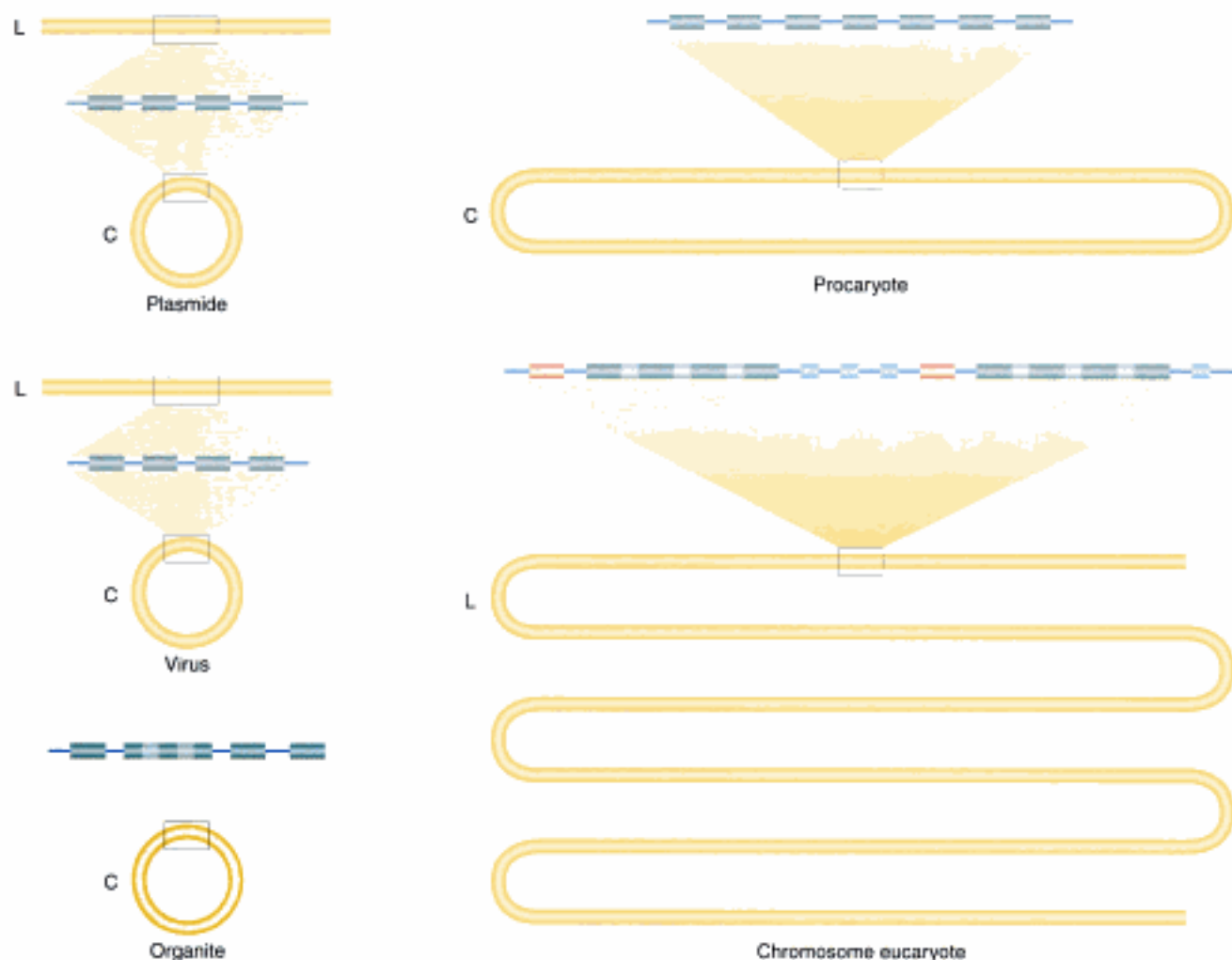


Figure 2-12 Représentation schématisée de la forme et de la taille approximative de différents types de génomes. L'agrandissement des segments montre les gènes en vert, les introns en vert clair et l'ADN répétitif en bleu ciel et en rouge.

La taille des chromosomes. Les chromosomes d'un même génome peuvent avoir des tailles très différentes. Dans le génome humain par exemple, le chromosome 1 (le plus grand) est quatre fois plus grand que le chromosome 21 (le plus petit), comme le montre le tableau 2-4. En étudiant les chromosomes d'une espèce, un cytogénéticien peut avoir des difficultés pour identifier les chromosomes d'après leur taille seule, mais il peut regrouper les chromosomes de tailles voisines. Par exemple, chez les humains, les chromosomes sont classés en sept groupes, de A (les plus grands) à G (les plus petits), comme le montre le tableau.

La position des centromères. Les fils moléculaires (les fibres du fuseau) qui tirent les chromosomes au cours de la division cellulaire s'attachent à une région spécialisée du chromosome appelée le **centromère**. La région centromérique ressemble généralement à une constriction ou à un étranglement en une position spécifique du chromo-

some. Cet étranglement divise le chromosome en deux « bras ». Le bras le plus court est appelé *p* et le bras le plus long, *q*. La position de l'étranglement définit le rapport de la longueur des deux bras du chromosome. Ce rapport est une caractéristique utile pour identifier les différents chromosomes (tableau 2-4). On peut classer les positions des centromères en **télocentrique** (à une extrémité), **acrocentrique** (près d'une extrémité), ou **métacentrique** (au milieu).

Les extrémités des chromosomes sont appelées **téломères**. Les télomères ne sont généralement pas différents à l'œil des autres parties du chromosome. Les télomères et les centromères ont des structures moléculaires uniques, qui sont essentielles au comportement normal des chromosomes.

La position des organisateurs nucléolaires. Les **nucléoles** sont des structures sphériques associées à des étranglements des chromosomes que l'on appelle des

TABLEAU 2-4
Chromosomes humains

Groupe	Nombre	Représentation schématique	Longueur relative*	Index centrométrique†
Grands chromosomes				
A	1		8,4	48 (M)
	2		8,0	39
	3		6,8	47 (M)
B	4		6,3	29
	5		6,1	29
Chromosomes moyens				
C	6		5,9	39
	7		5,4	39
	8		4,9	34
	9		4,8	35
	10		4,6	34
	11		4,6	40
	12		4,7	30
	13		3,7	17 (A)
D	14		3,6	19 (A)
	15		3,5	20 (A)
Petits chromosomes				
E	16		3,4	41
	17		3,3	34
	18		2,9	31
F	19		2,7	47 (M)
	20		2,6	45 (M)
G	21		1,9	31
	22		2,0	30
Chromosomes sexuels				
X			5,1 (groupe C)	40
Y			2,2 (groupe G)	27 (A)

* Pourcentage de la longueur totale cumulée d'un jeu haploïde de 22 autosomes.

† Longueur du bras court exprimée en pourcentage de la longueur totale du chromosome. Les quatre chromosomes les plus métacentriques sont indiqués par un (M) ; les quatre chromosomes les plus acrocentriques par un (A).

organiseurs nucléolaires (figure 2-13). Des organismes distincts possèdent des nombres différents de nucléoles, de un à un grand nombre par jeu de chromosomes. Les cellules diploïdes de nombreuses espèces possèdent une seule paire de nucléoles. Les organisateurs nucléolaires contiennent de nombreuses copies en tandem de gènes qui codent l'ARN ribosomal, un ARN non traduit qui entre dans la composition des ribosomes. L'ARN ribosomal est synthétisé au niveau des organisateurs nucléolaires, déposé dans les nucléoles, puis exporté dans le cytoplasme pour être incorporé dans les ri-

bosomes. Les positions des nucléoles, comme celles des centromères, sont des éléments très utiles pour l'analyse cytogénétique.

L'organisation des chromomères. Les **chromomères** ressemblent à des perles. Ce sont des renflements situés le long du chromosome au cours des premiers stades de la division nucléaire. Les positions des chromomères sont identiques dans tous les chromosomes homologues. Bien que les chromomères soient utiles comme marqueurs, leur nature moléculaire est inconnue.

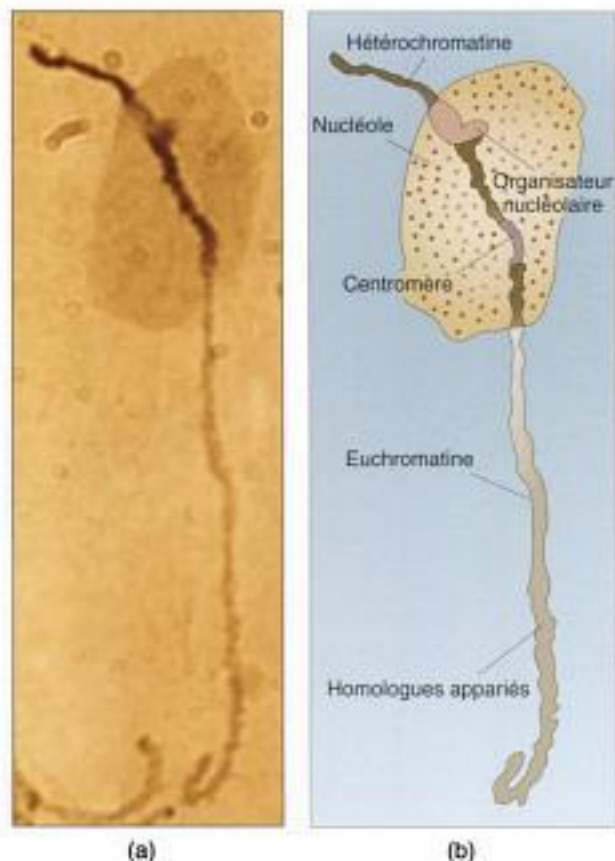


Figure 2-13 Le chromosome 2 de tomate au stade pachytène de la méiose. On voit le nucléole et l'organisateur nucléolaire : (a) photographie ; (b) interprétation. (Photographie Peter Moens. D'après *The genetic location of the centromere of chromosome 2 in the tomato*, P. Moens et L. Butler, Can. J. Genet. Cytol. 5 :364-370, 1963.)

L'organisation de l'hétérochromatine. Lorsque les chromosomes sont traités par des produits chimiques qui réagissent avec l'ADN, tels que le réactif de Feulgen, des régions distinctes avec des caractéristiques de marquage différentes apparaissent. Les régions fortement colorées sont appelées **hétérochromatine** et reflètent un haut degré de compacité. Les régions faiblement colorées sont appelées **euchromatine** et correspondent aux régions empaquetées de façon plus lâche. La plupart des gènes actifs se trouvent dans l'euchromatine. La figure 5-15 montre des exemples d'hétérochromatine chez la tomate.

L'organisation des bandes. Des procédures particulières de marquage des chromosomes ont révélé des séries de **bandes** (raies transversales) complexes chez de nombreux organismes différents. Les positions et les tailles des bandes sont hautement spécifiques des chromosomes. Elles représentent donc des éléments caractéristiques utiles. Il y a les bandes Q (produites par l'hydrochlorure de quinacrine), les bandes G (produites par la coloration au Giemsa) et les bandes R (les bandes qui ne sont pas colorées par le Giemsa). Les profils des bandes G dans deux chromosomes humains

et la position de plusieurs gènes sur ces chromosomes sont présentés dans la figure 2-14.

Un type assez particulier de bandes apparaît chez quelques organismes dont les chromosomes peuvent répliquer de nombreuses fois leur ADN sans en séparer les copies. Cela produit des chromosomes géants qui sont pratiquement des agrandissements des formes non répliquées. Grâce à cela, l'organisation des bandes normales des chromosomes devient facilement visible et peut servir d'élément de référence. Ces **chromosomes polytènes** (*polytène* signifie « contenant de nombreux fils ») se rencontrent dans des cellules hautement spécialisées des tubules de Malpighi, dans le rectum, les intestins, les extrémités des pattes et les glandes salivaires de diptères tels que les mouches communes, les moustiques et les mouches du vinaigre.

La mouche du vinaigre *Drosophila melanogaster* est un exemple très étudié. Chez cet insecte (un diploïde) les $2n$ chromosomes sont égaux à 8 et ces huit chromosomes (voir figure 2-15a) sont présents dans la plupart des cellules. Cependant, dans les cellules des organes qui contiennent les chromosomes polytènes, on observe des particularités intéressantes (figure 2-15b). Tout d'abord, il n'y a que quatre chromosomes polytènes par cellule (et non huit) car au cours du processus spécialisé de réplication, les membres de chaque paire d'homologues s'associent l'un à l'autre de façon inattendue. Deuxièmement, les quatre chromosomes polytènes sont reliés les uns aux autres au niveau d'une structure appelée le **chromocentre**, qui est une fusion de zones hétérochromatiques autour des centromères des quatre chromosomes. Le chromocentre des chromosomes de glande salivaire de la drosophile est présenté dans la figure 2-15b, où L et R désignent les bras, appelés de façon arbitraire, gauche (*left*) et droit (*right*).

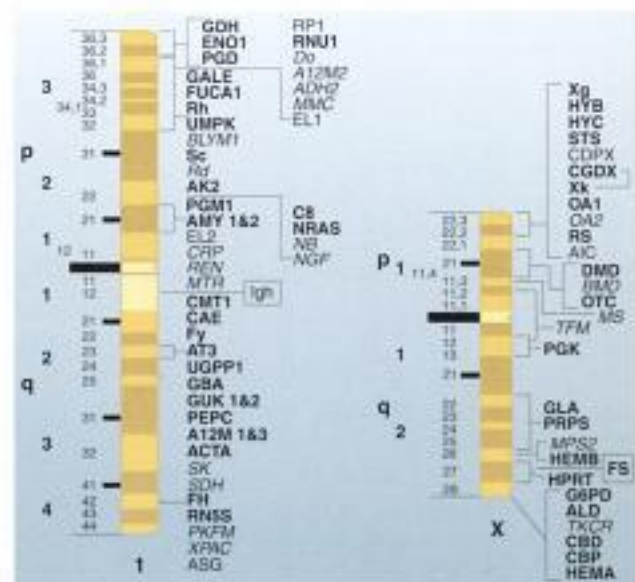
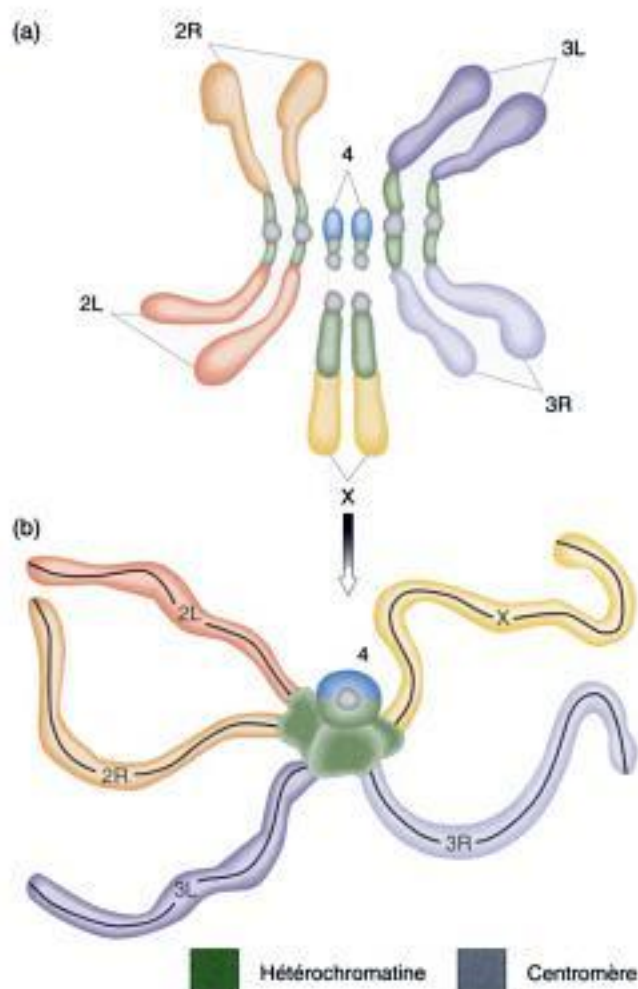


Figure 2-14 L'organisation des bandes sur les chromosomes humains 1 et X, indiquant la nomenclature des bandes et la position de certains gènes. (D'après V. A. McKusick, S. J. O'Brien, ed., *Genetic Maps*, vol.2. Cold Spring Harbor, N. Y. 1982.)



Le long d'un chromosome polytène apparaissent des bandes transversales (Figure 2-15c). Les bandes polytènes sont bien plus nombreuses que les bandes Q, G ou les bandes R. On les compte par centaines sur chaque chromosome. Les bandes diffèrent par leur largeur et leur aspect, de sorte que chaque chromosome a sa propre organisation de bandes.

Les études moléculaires ont montré que n'importe quelle région chromosomique de drosophile contient bien plus de gènes qu'il n'y a de bandes polytènes. Ces bandes ne représentent donc pas des gènes. L'interprétation des bandes Q, G et R n'est pas claire non plus chez les autres cellules eucaryotes. Si elles reflètent le degré de compacité de l'ADN, on ignore de quelle façon celui-ci est maintenu.

En utilisant la totalité des éléments caractéristiques des chromosomes, les cytogénéticiens peuvent identifier chacun des chromosomes de nombreuses espèces. La figure 2-16 montre une carte des caractéristiques chromosomiques du génome de maïs.

MESSAGE

Des caractéristiques telles que la taille, le rapport de la longueur des bras, l'hétérochromatine, le nombre et la position des chromomères, le nombre et la position des organisateurs nucléolaires et l'organisation des bandes permettent d'identifier chacun des chromosomes qui composent le jeu caractéristique de chaque espèce.

La structure tridimensionnelle des chromosomes nucléaires

Une cellule humaine contient environ 2 mètres d'ADN (1 m par jeu de chromosomes). Le corps humain est composé d'environ 10^{15} cellules et chacune est diploïde. Le



Figure 2-15 Les chromosomes polytènes forment un chromocentre dans une glande salivaire de drosophile. (a) Chromosomes en métaphase mitotique, avec les bras représentés par des teintes différentes. (b) L'hétérochromatine fusionne pour former le chromocentre. (c) Photographie de chromosomes polytènes (Tom Kaufman).

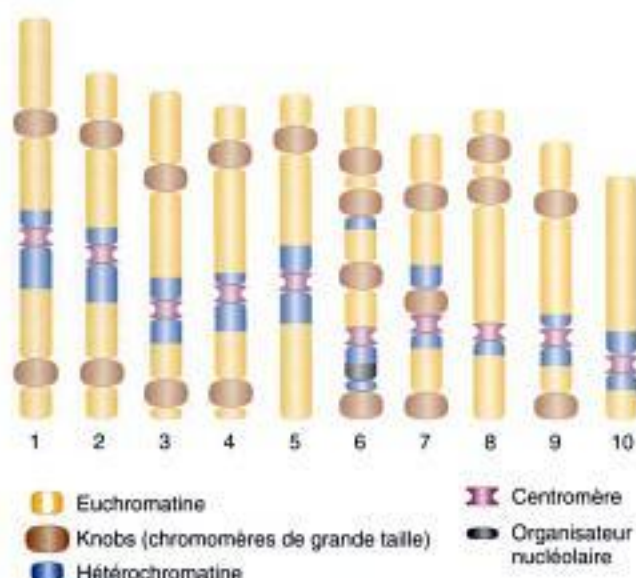


Figure 2-16 Les éléments caractéristiques qui différencient les chromosomes du maïs.

corps humain contient donc au total près de 2×10^{15} m d'ADN. On peut se faire une idée de cette longueur si importante de l'ADN en la comparant à la distance qui sépare la Terre du soleil, soit $1,5 \times 10^{11}$ m. L'ADN contenu dans notre corps pourrait donc relier la Terre au soleil et en revenir une centaine de fois. Grâce à cet exemple, on se rend compte que l'ADN des Eucaryotes est efficacement empaqueté. En fait cette compaction a lieu au niveau du noyau, dans lequel les 2 m d'ADN d'une cellule humaine sont empaquetés en 46 chromosomes, tous à l'intérieur d'un noyau de 0,006 mm de diamètre. Comment de si longues molécules peuvent-elles être empaquetées en ces structures qui ressemblent à des bâtonnets, visibles sous microscope photonique et que nous appelons des chromosomes ? Pour répondre à cette question, nous devons comprendre la structure tridimensionnelle des chromosomes eucaryotes.

Si les cellules eucaryotes sont cassées et que l'on examine le contenu de leur noyau au microscope électronique, les chromosomes apparaissent comme des masses de spaghetti d'un diamètre proche de 30 nm. On en voit quelques exemples dans la figure 2-17. Si l'on étudie avec attention ces chromosomes, il devient évident qu'il n'y a pas d'extrémité saillant de la masse de fibres. Ceci suggère que chaque chromosome est une fibre longue et fine, repliée d'une façon ou d'une autre. Si la fibre correspond à une molécule d'ADN, on arrive alors à l'idée que chaque chromosome est une molécule d'ADN fortement repliée. La microscopie électronique (figure 2-18) et des mesures de changements de conformation de l'ADN de drosophile en solution soutiennent cette idée.

Les arguments les plus forts du fait qu'un chromosome est constitué d'une seule longue molécule d'ADN viennent sans doute d'une technique appelée *électrophorèse en champs alternés*. Si l'on extrait l'ADN d'un grand nombre de cellules, qu'on le purifie avec soin pour éviter

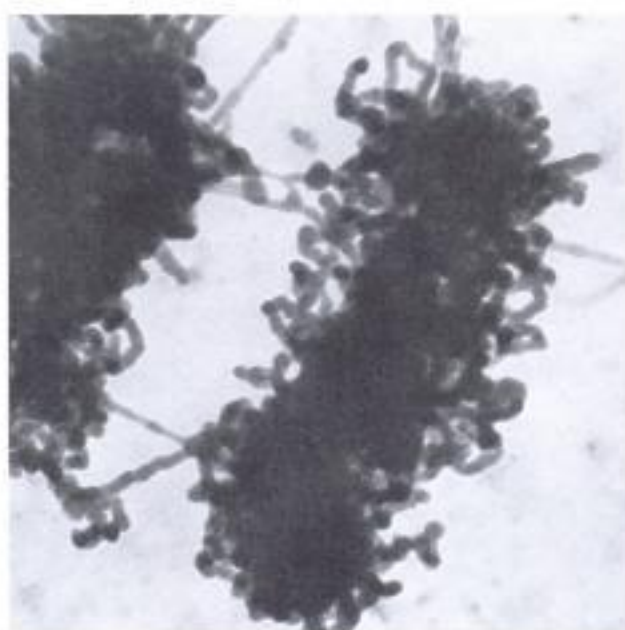


Figure 2-17 Micrographie électronique de chromosomes métaphasiques d'abeille. Les chromosomes apparaissent chacun composés d'une fibre continue de 30 nm de large. (D'après E. J. DuPraw, *Cell and Molecular Biology*. Copyright ©1968 par Academic Press.)

de casser les molécules, et qu'on place celles-ci sur une matrice gélatineuse sous l'influence de champs électriques puissants, croisés et alternés, les molécules d'ADN de chaque chromosome se déplacent dans le gel à une vitesse proportionnelle à leur taille. Tous les ADN d'un type de chromosome finissent leur course à la même position (que l'on appelle une *bande*), de sorte que si le gel est coloré avec une substance qui se fixe à l'ADN, le nombre de bandes d'ADN est toujours égal au nombre de chromosomes, que la cellule soit haploïde ou diploïde.



Figure 2-18 Micrographie électronique reconstituée d'une molécule d'ADN de drosophile. La longueur totale est de 1,5 cm et on pense qu'elle correspond à un chromosome. (D'après R. Kavenoff, L. C. Klotz et B. H. Zimm, *Cold Spring Harbor Symp. Quant. Biol.* 38, 1974, 4.)

MESSAGE

Chaque chromosome eucaryote contient une seule longue molécule repliée d'ADN.

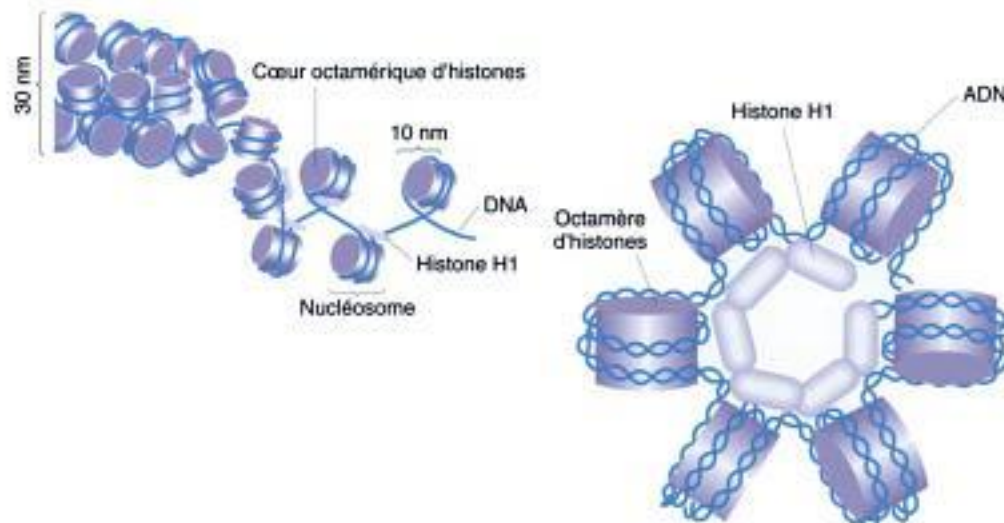
Les protéines histones. Quels sont les mécanismes précis qui permettent d'empaqueter l'ADN en chromosomes ? Comment le très long fil d'ADN est-il converti en une structure dense, sinueuse et relativement épaisse, qui ressemble à un bâtonnet, et qui est visible au cours de la division cellulaire ? Tout le matériel qui constitue les chromosomes est désigné par le terme général de **chromatine**. Il s'agit d'ADN et de protéines. Si la chromatine est extraite et traitée avec des concentrations salines différentes, on observe des degrés différents de compaction ou de condensation sous le microscope élec-

tronique. À faible concentration saline, on voit une structure d'environ 10 nm de diamètre, qui ressemble à un collier de perles. Le fil entre les perles du collier peut être digéré par l'enzyme ADNase. On peut donc en déduire que le fil est de l'ADN. On appelle les perles du collier, des **nucléosomes**. On peut montrer qu'il s'agit de complexes d'ADN et de protéines chromosomiques spéciales que l'on appelle des **histones**. La structure des histones est remarquablement conservée entre les organismes eucaryotes et tous les nucléosomes contiennent un octamère constitué de deux unités de chacune des histones H2A, H2B, H3 et H4. L'ADN est enroulé deux fois autour de l'octamère comme le montre la figure 2-19a. Lorsque les concentrations de sel sont plus élevées, le collier de perles adopte progressivement l'aspect d'un **solénoïde** (figure 2-19b). Ce solénoïde, produit *in vitro*, a un diamètre de 30 nm et correspond probablement aux structures en spaghetti visibles *in vivo* comme dans la figure 2-17. Le solénoïde est stabilisé par une autre histone, H1, qui se trouve au centre de la structure, comme le montre la figure 2-19b.



(a)

Figure 2-19 (a) Modèle d'un nucléosome montrant l'ADN enroulé deux fois autour d'un octamère d'histones. (Alan Wolffe et Van Moudrianakis). (b) Deux représentations d'un modèle du solénoïde de 30 nm de diamètre, montrant les octamères d'histones sous la forme de disques violets. (A gauche) Vue latérale partiellement déroulée. (A droite) Vue de l'enroulement complet. L'histone H1 supplémentaire est représentée au centre de l'enroulement, où elle joue probablement le rôle de stabilisateur. Lorsqu'on augmente les concentrations salines, les nucléosomes se rapprochent pour former un solénoïde avec six nucléosomes par tour. (D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipursky, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997.)



(b)

Nous voyons donc que pour atteindre le premier niveau de compaction, l'ADN s'enroule autour des histones, qui jouent en quelque sorte le rôle de bobines. Le degré d'enroulement supérieur aboutit à la conformation en solénoïde. Cependant, il faut encore un degré d'enroulement supplémentaire pour convertir les solénoïdes en cette structure tridimensionnelle que nous appelons le chromosome.

Les niveaux supérieurs d'enroulement. De nombreuses études cytogénétiques montrent que les chromosomes sont enroulés. La figure 2-20 présente un bon exemple tiré du noyau d'un protozoaire. Alors que le diamètre du solénoïde est de 30 nm, le diamètre de ces « super-tours » est le même que le diamètre du chromosome au cours de la division cellulaire, soit généralement de l'ordre de 700 nm. Qu'est-ce qui produit ce superenroulement ? L'observation de chromosomes dont on a retiré chimiquement les histones apporte un élément de réponse. Après un tel traitement, les chromosomes apparaissent avec une partie centrale, appelée **armature**, formée de protéines non histones présentant un marquage dense, que l'on voit dans la figure 2-21. De cette armature protéique partent des boucles d'ADN. On pense que les fibres qui forment ces boucles sont les solénoïdes. À un grossissement supérieur, sur des micrographies électroniques, il est clair que chaque boucle d'ADN commence et finit au niveau de l'armature. On a découvert que l'armature centrale était en grande partie composée de l'enzyme topoisomérase II, qui est capable de faire passer un brin d'ADN au travers d'un autre, un peu comme des anneaux de magicien. À l'évidence, l'armature centrale



Figure 2-20 Dessins de chromosomes d'un protozoaire, montrant l'enroulement et le superenroulement visibles au microscope photonique. On peut voir deux grands chromosomes : l'un en orange et l'autre en jaune ; les chromosomes sont dupliqués car la division de cellules sexuelles est sur le point de se produire. (D'après L. R. Cleveland, « The whole life of chromosomes and their coiling systems », *Trans. Am. Philosophical Soc.* 39, 1949, 1.)



Figure 2-21 Micrographie électronique d'un chromosome en métaphase issu d'une cellule humaine en culture. Remarquez l'armature, ou cœur, à partir de laquelle les brins d'ADN s'étendent vers l'extérieur. Aucune extrémité n'est visible à la périphérie. Même à des grossissements supérieurs, il est clair que chaque boucle commence et finit près de la même région de l'armature. (D'après W. R. Baumbach et K. W. Adolph, *Cold Spring Harbor Symp. Quant. Biol.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., 1977.)

organise le vaste écheveau d'ADN au cours de la réplication, un moment où les deux brins d'ADN doivent se dérouler pour permettre une nouvelle synthèse. Quiconque a essayé de séparer les deux brins d'un morceau de corde s'est rendu compte des difficultés inhérentes à un tel processus.

Interrogeons-nous à nouveau sur la façon dont le superenroulement du chromosome est produit. Les arguments les plus forts suggèrent que le solénoïde est arrangé en boucles qui partent de l'armature centrale, qui a elle-même la forme d'une spirale. Nous en voyons une représentation générale dans la figure 2-22, qui montre un chromosome enroulé lâchement dans une cellule en interphase, et un chromosome enroulé de façon plus étroite dans une cellule en division. Les boucles sont fixées à l'armature au niveau de régions spéciales le long de l'ADN, appelées **régions de liaison à l'armature**, ou SAR (pour *scaffold attachment regions*). Certaines SAR cartographiées chez la drosophile sont présentées dans la figure 2-23.

Comme nous l'avons vu dans la figure 2-22, les chromosomes sont bien moins enroulés (sans qu'il s'agisse

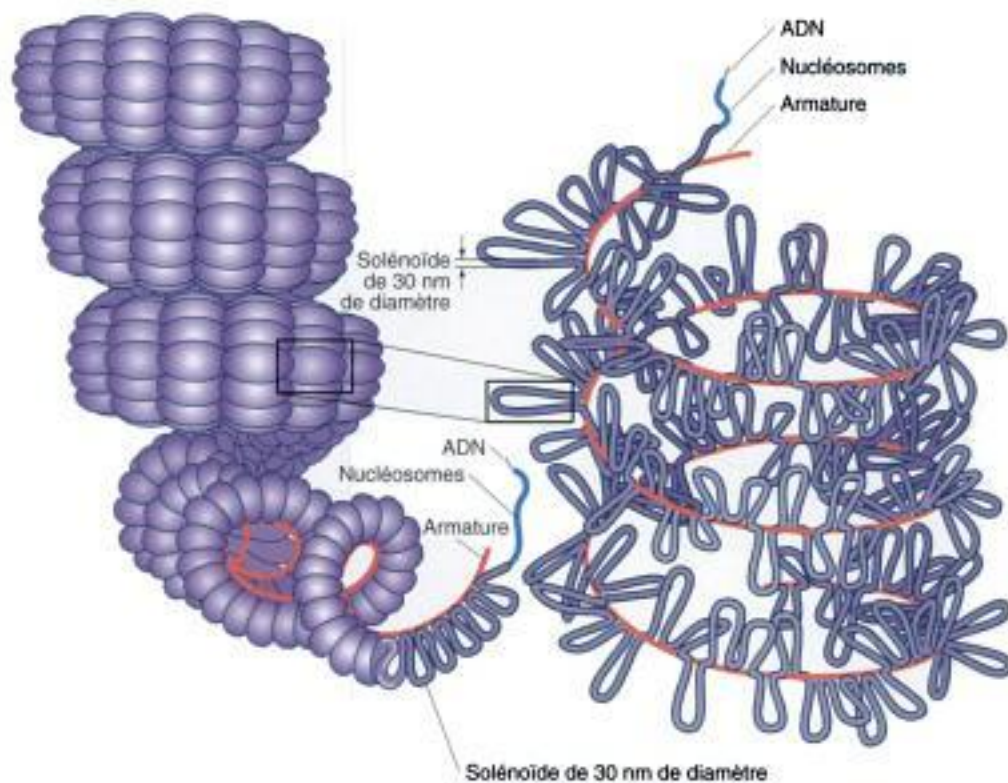


Figure 2-22 Modèle de la structure d'un chromosome. Sur la droite, on voit un supertour plus relâché, lorsque la cellule est en interphase. Sur la gauche est présenté un enroulement plus dense, correspondant à un chromosome en cours de division cellulaire : là, les boucles sont empaquetées si étroitement que seules leurs extrémités sont visibles. Au niveau des extrémités libres, les solénoïdes sont montrés déroulés pour donner une idée approximative de l'échelle relative.

d'ADN nu, c'est-à-dire sans protéine) dans un noyau en interphase qu'au cours de la division cellulaire. C'est pourquoi ils sont quasiment invisibles au microscope photonique. Cet état plus relâché est probablement nécessaire pour l'expression de la plupart des gènes.

MESSAGE

Il existe des niveaux successifs d'empaquetage des chromosomes

1. L'ADN s'enroule autour des bobines d'histones.
2. L'enchaînement des nucléosomes s'enroule en un solénoïde.
3. Le solénoïde forme des boucles et ces boucles se fixent à l'armature centrale.
4. L'armature et les boucles s'organisent en un super-enroulement géant.

Les représentations des chromosomes. Nous avons vu que les chromosomes eucaryotes sont des structures complexes, et en conséquence, ce qui n'a rien d'étonnant, ils sont difficiles à schématiser. La figure 2-24 montre trois représentations stylisées de chromosomes haploïdes et diploïdes dans des cellules en interphase. L'accent est mis d'abord sur la structure enroulée, puis sur le contenu d'ADN linéaire, et enfin sur la longueur des chromosomes et le rapport de la longueur relative des bras. L'une ou l'autre de ces représentations est utilisée en fonction des

propriétés spécifiques des chromosomes qui sont discutées, mais il faut souligner qu'aucune de ces représentations n'est parfaite. Les schémas présentés dans la figure 2-24, à la page suivante, sont des représentations des chromosomes dans une cellule en interphase, et comme nous l'avons écrit précédemment, à ce stade, les chromosomes ne sont même pas visibles au microscope. Il est important de remarquer que dans l'état diploïde présenté, les chromosomes homologues constituent une paire, mais qu'ils ne sont généralement pas dans une configuration appariée. En fait, les chromosomes homologues ne s'apparient physiquement qu'au cours du type spécialisé de division cellulaire qui a lieu lors de la production de cellules sexuelles (méiose), comme nous le verrons au chapitre 4.

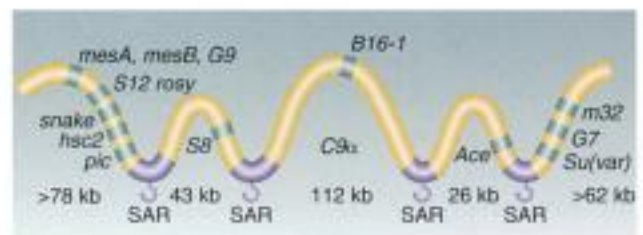


Figure 2-23 Certains domaines en boucle ont été cartographiés chez la drosophile. On y voit les locus de certains gènes et des régions de liaison à l'armature (SAR).

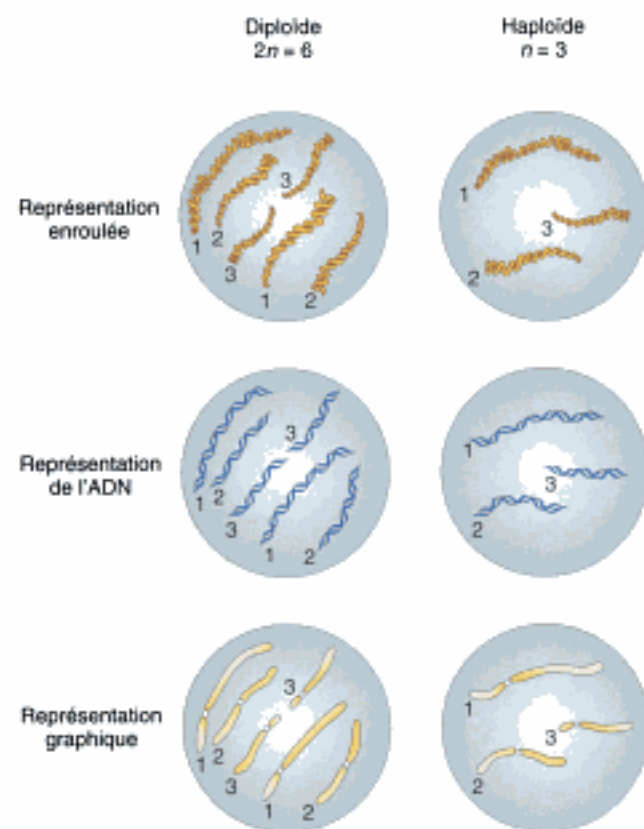


Figure 2-24 Trois façons schématiques de représenter les chromosomes dans des cellules en interphase. La première représentation met l'accent sur la structure enroulée des chromosomes. La deuxième souligne le fait que le composant génétique d'un chromosome est de l'ADN. La troisième montre les chromosomes sous la forme de « vers de terre » avec des centromères représentés par des ronds blancs. Aucune représentation n'est parfaitement juste mais chacune est utilisée en génétique.

RÉSUMÉ

Le phénotype d'un organisme est déterminé par le phénotype de ses cellules. C'est pourquoi, pour comprendre la façon dont les gènes déterminent les caractéristiques d'une espèce, il est nécessaire de comprendre la nature des génomes au niveau cellulaire.

Presque tous les génomes sont constitués d'ADN, une molécule en forme de double hélice composée de chaînes nucléotidiques entrelacées et antiparallèles. On retrouve toujours quatre nucléotides : A, T, G et C. Un A dans un brin est toujours apparié avec un T dans l'autre brin, en raison de la formation de liaisons hydrogènes complémentaires. Pour les mêmes raisons, un G dans un brin s'apparie toujours avec un C dans l'autre brin. La séquence nucléotidique de l'ADN représente un ensemble d'informations codées génétiquement dont la finalité est de spécifier la forme biologique.

Les Procaryotes (bactéries) ont un chromosome circulaire présent en une ou plusieurs copies. Les Eucaryotes ont soit un, soit deux exemplaires de leur génome

nucléaire, selon qu'ils sont haploïdes ou diploïdes. Le génome nucléaire eucaryote est constitué de chromosomes linéaires, en petit ou en grand nombre. Les cellules eucaryotes contiennent également des chromosomes plus petits dans les mitochondries et les chloroplastes. Il s'agit de chromosomes circulaires présents en de nombreuses copies. Les virus, bien qu'il ne s'agisse pas d'organismes, possèdent également de petits génomes. Un gène est une unité fonctionnelle située dans l'ADN d'un chromosome. On peut définir un gène comme une région d'ADN susceptible de donner naissance à un transcrit d'ARN fonctionnel, au bon moment, au bon endroit et dans des quantités appropriées pour permettre un développement normal. Chaque chromosome eucaryote contient une molécule d'ADN. Les chromosomes eucaryotes sont empaquetés efficacement par enroulement et superenroulement en utilisant des structures semblables à des bobines, constituées d'histones et d'une armature centrale.

DIAGRAMME CONCEPTUEL

Dessinez un diagramme conceptuel mettant en relation le maximum de termes parmi ceux qui sont proposés. Remarquez que ces termes ne sont pas écrits dans un ordre particulier.

Double hélice d'ADN / génome / noyau / homologues / ploidie / n / haploïde / diploïde / chromosome

PROBLÈME RÉSOLU

1. On a établi qu'une molécule d'ADN en double hélice avait un rapport bases puriques : bases pyrimidiques de 1 : 4 dans une chaîne nucléotidique. Quel doit être le rapport bases puriques : bases pyrimidiques dans l'autre brin ?

SOLUTION

Quelles sont les connaissances nécessaires sur l'ADN ici ? Nous savons d'après la structure générale de l'ADN, qu'une base purique dans un brin s'apparie toujours avec une base pyrimidique dans l'autre brin (la base purique A s'apparie avec la base pyrimidique T et la base purique G s'apparie avec la base pyrimidique C). Dans le brin en question, les bases puriques occupent $1/(1+4)$, soit $1/5$

des sites d'un brin. Dans l'autre brin, $1/5$ des sites doivent donc être des bases pyrimidiques. De la même façon, $4/5$ des sites dans l'autre brin doivent être des bases puriques. Donc, le rapport bases puriques : bases pyrimidiques dans l'autre brin sera de $4/5 : 1/5$, soit $4 : 1$. (Remarquez qu'il s'agit de l'inverse du rapport de l'autre brin.)

PROBLÈME RÉSOLU

2. Le génome humain contient 3 000 000 kb et on estime qu'il comporte 100 000 gènes. D'après ces données, si l'on fait l'hypothèse simple que la taille moyenne de la région codante d'un gène (la partie qui code une protéine) est de 2 kb,

- a. Que peut-on conclure sur l'espace moyen sé-

parant les milieux de deux gènes consécutifs sur l'ADN chromosomique ?

- b. Que peut-on conclure sur le pourcentage du génome qui code des protéines ?
c. Que contiennent les régions non codantes ?

SOLUTION

- a. L'espacement moyen entre les milieux des gènes doit être de 3 000 000 divisé par 100 000, soit 30 kb.
b. Puisque le segment codant une protéine est seulement de 2 kb, il est clair que seul $1/15$, soit environ 6 % du génome code effectivement des protéines.

- c. Les autres séquences doivent contenir des régions régulatrices adjacentes aux gènes, ainsi que de l'ADN répétitif et de l'ADN non répétitif situé entre les régions codant des protéines et les introns présents à l'intérieur des régions codantes.

PROBLÈMES

1. Dessinez un schéma qui montre les relations hiérarchiques entre les termes *eucaryote*, *procaryote*, *haploïde*, *diploïde*, *génome nucléaire*, *génome chloroplastique*, *génome mitochondrial*, *plasmides*.
2. Dans une seule cellule de feuille d'un végétal, combien de copies y a-t-il du génome nucléaire, du génome chloroplastique et du génome mitochondrial (choisissez entre « beaucoup », « deux » ou « une ») ?
3. Que signifie le terme *homologue* ?
4. Les 3 milliards de paires de bases du génome humain ont été comparées à l'information contenue

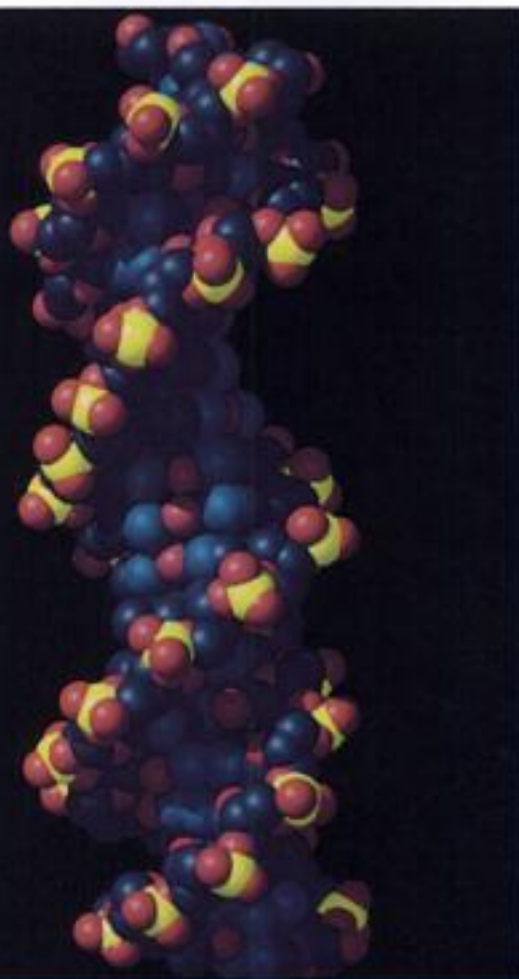
dans 1000 livres. Si la longueur totale de l'ADN du génome est de 1 m, de quelle taille serait un livre en termes de

- a. Paires de bases
b. Longueur d'ADN
5. a. Parmi les organismes suivants, lesquels possèdent un génome mitochondrial : un poisson, une mousse, un palmier, une levure de boulanger, une bactérie ?
b. Parmi les organismes suivants, lesquels possèdent un génome chloroplastique : une diatomée, un serpent, un champignon, la bactérie des intestins *E. coli*, du gui ?

6. Dressez la liste des forces qui maintiennent la double hélice d'ADN sous la forme d'une unité stable.
7. Si la thymine constitue 15 pour cent des bases d'un échantillon d'ADN donné, quel sera le pourcentage de la cytosine ?
8. Si le contenu de G+C d'un échantillon d'ADN est de 48 pour cent, quel sera le pourcentage de chacun des quatre nucléotides ?
9. Dessinez un schéma simple de l'ADN qui indique clairement où sont situées les extrémités 3' et 5'.
10. Chaque cellule du corps humain (à l'exception des gonades) contient 46 chromosomes. Quel est le niveau de ploïdie et le nombre haploïde de chromosomes ?
11. Combien de molécules d'ADN y a-t-il dans une cellule du corps humain ?
12. Quels sont les différents niveaux d'organisation de l'ADN qui aboutissent aux structures visibles que l'on appelle des chromosomes ?
13. Un certain segment d'ADN a la séquence nucléotidique suivante dans l'un de ses brins :
 5'ATTGGCTCT3'
 Quelle doit être la séquence sur l'autre brin (précisez ses extrémités 5' et 3') ?
14. Dans une double hélice d'ADN normale, est-il vrai que
 - a. A+C est toujours égal à G+T ?
 - b. A+G est toujours égal à C+T ?
15. Les éléments les plus courants chez les organismes vivants sont le carbone, l'hydrogène, l'oxygène, l'azote, le phosphore et le soufre. Quel est celui que l'on ne rencontre pas dans l'ADN ?
16. Qu'est-ce qu'un gène ? Citez des problèmes posés par votre définition.
17. Le génome de la bactérie *Hemophilus influenzae* a une taille de 1830 kb et son séquençage a permis de savoir qu'il contient 1703 gènes.
 - a. Quel est l'espacement moyen entre les milieux de deux gènes consécutifs ?
 - b. Si la taille moyenne d'une région codante est estimée à 1 kb, quelle est la distance moyenne entre les gènes ?
 - c. Qu'y a-t-il dans cette région ?
 - d. À quel pourcentage du génome s'attend-on pour les introns ?
- e. Quel est l'espacement moyen entre les régions codantes ?
18. Le gène de l'albumine humaine s'étend sur une région chromosomique de 25 000 paires de nucléotides (25 kb) du début de la séquence codant la protéine à la fin de cette même séquence, mais l'ARN messager de cette protéine n'est long que de 2.1 kb. À votre avis, qu'est-ce qui explique cette différence importante ?
19. On extrait de l'ADN des cellules de *Neurospora*, un champignon haploïde ($n = 7$), de petit pois, un végétal diploïde ($2n = 14$), de la mouche commune, un animal diploïde ($2n = 12$). Si ces ADN sont séparés au moyen d'une électrophorèse en champs alternés, combien de bandes produira chacune des trois espèces ?
20. Imaginez une formule qui mette en rapport la taille d'un ARNm, la taille du gène correspondant, le nombre d'introns et la taille moyenne de ces introns.
21. Des progrès dans le séquençage du génome ont permis d'estimer qu'il existe une différence de 1,7 pour cent entre le génome humain et celui du chimpanzé. En supposant que les deux génomes ont approximativement la même taille, combien de différences de nucléotides y a-t-il entre ces deux espèces ?
22. En utilisant des phrases, écrivez
 - a. La différence entre les bases puriques et les bases pyrimidiques dans l'ADN
 - b. La différence entre l'adénine et la guanine
 - c. La différence entre la cytosine et la thymine
23. Pourquoi les extrémités 5' et 3' sont-elles appelées ainsi ?
24. Dessinez les segments suivants d'ADN de la même façon que dans la figure 2-2b :
 5'ATGCT3'
 3'TACGA5'
 3'ATGCT5'
 5'TACGA3'
25. Combien de bases y a-t-il dans 2 kb d'ADN ?
26. En utilisant les figures de ce chapitre, estimez le nombre de paires de bases par tour complet de double hélice d'ADN.
27. Chez un insecte donné, $2n = 18$. Si les chromosomes polytènes se forment exactement comme chez la drosophile, combien de chromosomes polytènes trouvera-t-on par cellule dans les glandes salivaires ?

3

LA FONCTION DES GÈNES



Modélisation informatique de l'ADN.
(J. Newdell, Computer Graphics Laboratory, University of California, San Francisco. Copyright © Regents, University of California.)

POUR L'ESSENTIEL

Les fonctions de l'ADN et de l'ARN découlent de la complémentarité entre les acides nucléiques simple-brin et de la fixation de différents types de protéines sur les acides nucléiques.

La plupart des gènes codent des protéines ; quelques-uns codent des ARN non traduits qui sont des composants fonctionnels du système de traitement de l'information.

Dans les gènes codant des protéines, l'ADN est transcrit en un ARN, qui est à son tour traduit en une chaîne polypeptidique.

Un seul brin d'ADN est utilisé comme matrice pour la transcription d'un gène donné.

Lors de la polymérisation de l'ARN, les nouveaux nucléotides sont toujours ajoutés à l'extrémité 3' de la molécule en cours d'élongation.

La fonction des protéines est déterminée par la séquence primaire d'acides aminés, qui est spécifiée par la séquence nucléotidique.

La traduction de l'ARNm se déroule trois nucléotides (un codon) à la fois ; la spécificité naît de la complémentarité de la liaison entre le codon et l'anticodon d'un ARNt portant l'acide aminé adéquat.

Les acides aminés sont ajoutés du côté de l'extrémité carboxyle de la chaîne polypeptidique en croissance.

La fonction d'une protéine est déterminée par sa forme, sa taille et ses propriétés de liaison, toutes trois spécifiées par la séquence d'acides aminés.

Des changements dans la séquence nucléotidique de l'ADN peuvent conduire à des changements dans la séquence d'acides aminés, qui entraînent souvent un dysfonctionnement de la protéine.

Le dysfonctionnement d'une protéine est le plus souvent récessif, mais peut être parfois dominant.

La mucoviscidose est une maladie humaine due à la présence d'un allèle déficient d'un seul gène. Cette maladie est si grave que dans le passé, elle tuait généralement les malades dans leur enfance. Les symptômes affectent de nombreuses parties du corps (figure 3-1). Un des symptômes principaux est une sueur excessivement salée, mais la cause la plus fréquente du décès est un blocage des voies aériennes par un type de mucus inhabituellement épais. Comment est-il possible qu'un défaut portant sur seulement l'un des 100 000 gènes du génome humain puisse entraîner la mort ? À l'évidence, ce gène assure une fonction essentielle au bon déroulement de la vie humaine. En fait, la fonction des gènes est un élément clé pour comprendre cette maladie héréditaire et les milliers d'autres maladies qui affectent la santé et peuvent conduire à une mort prématurée. Dans ce chapitre, nous verrons les mécanismes du fonctionnement des gènes dans les organismes en général, puis nous reviendrons sur la mucoviscidose pour en examiner le gène responsable, sa fonction normale et ce que nous savons actuellement sur la façon dont l'allèle déficient provoque ces multiples symptômes.

LES GÈNES ET LES ARN

Comme nous l'avons vu précédemment, les produits initiaux de tous les gènes sont les acides ribonucléiques (ARN). L'ARN est fabriqué au cours d'un processus qui copie la séquence nucléotidique de l'ADN. Puisque ce processus rappelle la transcription (copie) de documents écrits, la synthèse de l'ARN est appelée **transcription**. On dit que l'ADN est transcrit en ARN et on appelle l'ARN un **transcrit**. Comme nous le verrons, le processus par lequel les ARN sont transcrits à partir des gènes présente certaines ressemblances avec la façon dont un nouvel ADN est synthétisé au cours de la réplication de l'ADN, mais il existe aussi des différences importantes.

Une façon d'envisager les rôles biologiques distincts de l'ADN et de l'ARN est de considérer l'ADN (c'est-à-dire le génome) comme le manuel d'instructions nécessaires pour produire tous les ARN dont la cellule a besoin, tandis que l'ARN n'est que la copie, qui peut être perdue, des parties du manuel concernant la réalisation d'une tâche donnée. Nous avons examiné l'ADN auparavant. Pour comprendre les rôles joués par l'ARN dans une cellule, intéressons-nous d'abord à ses caractéristiques chimiques générales.

Les propriétés de l'ARN

Bien que l'ARN et l'ADN soient tous deux des acides nucléiques, l'ARN est différent sur plusieurs points importants :

1. L'ARN est une chaîne nucléotidique simple-brin et non une double hélice. L'une des conséquences en est que l'ARN peut adopter des formes tridimensionnelles bien plus complexes que les molécules d'ADN double-brin. Nous reviendrons plus en détail sur cet aspect dans ce chapitre.

Organes touchés par la mucoviscidose

Le défaut génétique à l'origine de la mucoviscidose perturbe le fonctionnement de plusieurs organes en obstruant les conduits et autres canaux, en général par un mucus épais et visqueux ou par d'autres sécrétions.

Voies aériennes

L'obstruction et l'infection des voies bronchiques empêche la respiration. Les infections détruisent progressivement les poumons. La maladie des poumons est responsable de la plupart des décès dus à la mucoviscidose.

Foie

L'obstruction des petits canaux véhiculant la bile empêche la digestion et perturbe le fonctionnement du foie chez environ 5 pour cent des malades.

Pancréas

L'occlusion des canaux empêche le pancréas de sécréter certaines enzymes digestives essentielles aux intestins chez environ 85 pour cent des malades. Des diabètes peuvent également apparaître.

Intestin grêle

L'obstruction de l'intestin par des selles épaisses nécessite une intervention chirurgicale chez près de 10 pour cent des nouveau-nés.

Tractus génital

L'absence de fins canaux tels que les canaux déférents, rend stériles 95 pour cent des hommes atteints de mucoviscidose. Parfois, les femmes malades sont stériles en raison de la présence d'un bouchon dense de mucus qui empêche les spermatozoïdes de pénétrer dans l'utérus.

Peau

Un dysfonctionnement des glandes sudoripares entraîne la présence d'un excès de sel (chlorure de sodium) dans la sueur. La mesure du chlorure dans la sueur est un élément important du diagnostic.

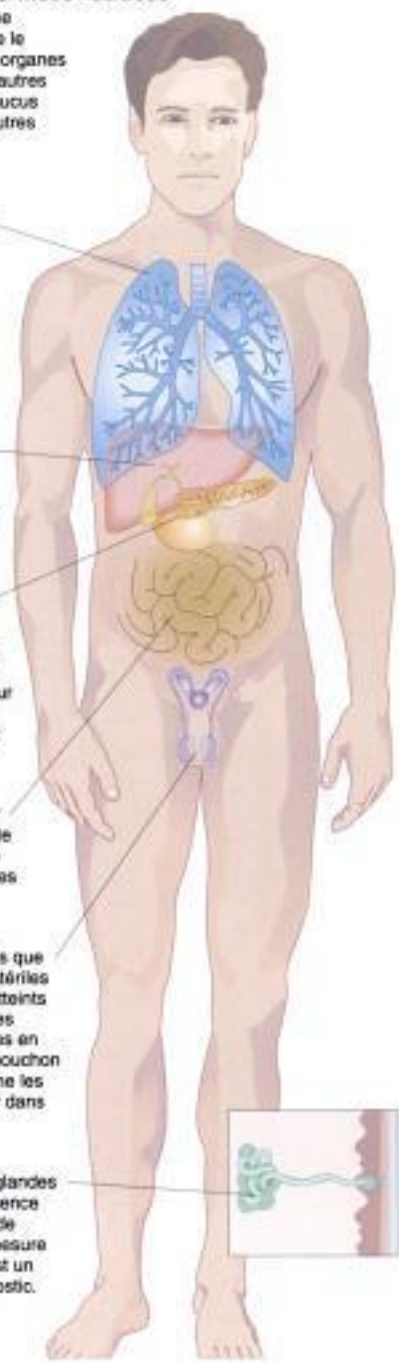
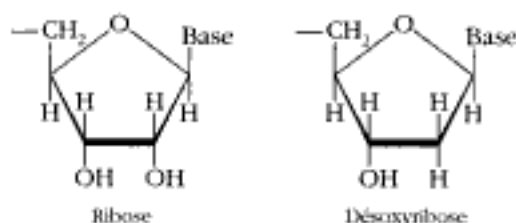
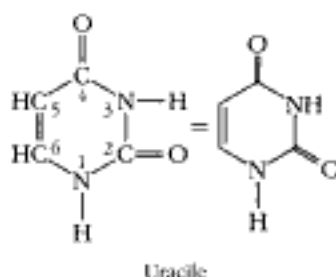


Figure 3-1 Les organes touchés par la mucoviscidose.

2. Le sucre contenu dans les nucléotides de l'ARN est le **ribose** et non le désoxyribose. Comme leurs noms le suggèrent, les deux sucres diffèrent par la présence ou l'absence d'un atome d'oxygène. Comme pour les brins individuels de l'ADN, l'ARN possède un squelette phosphate-ribose, avec une base liée de façon covalente à la position 1' de chaque ribose.



3. Les nucléotides de l'ARN portent les bases adénine, guanine et cytosine, mais la base pyrimidique **uracile** (abrégiée en U) remplace la thymine.



Toutefois, l'uracile forme des liaisons hydrogène avec l'adénine, exactement comme la thymine.

Les classes d'ARN

Les ARN peuvent être regroupés en deux grandes classes. Certains ARN sont des intermédiaires dans le processus de décodage des gènes en chaînes polypeptidiques. Nous appellerons ces ARN, des ARN « informationnels ». Dans l'autre classe, c'est l'ARN lui-même qui est le produit final fonctionnel. Nous parlerons alors d'ARN « fonctionnels ».

Les ARN informationnels. Pour la plupart des gènes, l'ARN est seulement un intermédiaire dans la synthèse du produit fonctionnel final, qui est une protéine. L'ARN informationnel de cette majorité de gènes est toujours un ARN messager (ARNm).

Chez les Procaryotes, le transcrit, qui est synthétisé sans modification à partir de l'ADN (le transcrit primaire) est l'ARNm. Chez les Eucaryotes en revanche, le transcrit primaire subit une maturation. Il s'agit de modifications au niveau de ses extrémités 5' et 3' et du retrait de certains éléments du transcrit primaire (introns). À la fin de cette maturation du pré-ARNm, un ARNm est produit.

Nous verrons ces étapes de la production des ARNm plus loin dans ce chapitre.

La séquence des nucléotides dans l'ARNm est convertie en une séquence d'acides aminés appartenant à une chaîne polypeptidique, par un processus appelé **traduction**. Dans ce cas, le mot *traduction* est utilisé dans le sens de traduction d'une langue étrangère : la cellule dispose d'un moyen pour traduire le langage de l'ARN dans le langage des polypeptides. Les protéines sont constituées d'une ou plusieurs chaînes polypeptidiques.

Les ARN fonctionnels. Grâce à la connaissance approfondie de la biologie cellulaire, nous avons compris que les ARN fonctionnels remplissent de nombreux rôles. Une fois encore, il est important de souligner que l'action de ces ARN fonctionnels se situe exclusivement au niveau des ARN ; ils ne sont jamais traduits en polypeptides. Chaque classe d'ARN fonctionnel est codée par un nombre relativement faible de gènes (de quelques dizaines à quelques centaines).

Les principales classes d'ARN fonctionnels interviennent dans différentes étapes du traitement informationnel de l'ADN en protéine. Deux de ces classes d'ARN fonctionnels se rencontrent chez tous les organismes : les ARN de transfert et les ARN ribosomaux.

Les molécules d'**ARN de transfert (ARNt)** agissent comme des transporteurs qui apportent des acides aminés à proximité de l'ARNm au cours du processus de traduction (synthèse protéique). Les ARNt sont des composants généraux de la machinerie de traduction. Ils peuvent apporter des acides aminés à l'ARN de tout gène codant une protéine.

Les **ARN ribosomaux (ARNr)** sont des composants des ribosomes. Ceux-ci sont de gros ensembles macromoléculaires qui servent de guides pour coordonner l'assemblage de la chaîne d'acides aminés d'une protéine. Les ribosomes sont constitués de plusieurs types d'ARNr et d'environ 100 protéines différentes. Comme dans le cas de l'ARNt, les ARNr sont les composants traductionnels généraux qui peuvent être utilisés pour traduire l'ARNm de n'importe quel gène codant des protéines.

Deux autres classes d'ARN fonctionnels impliqués dans le traitement de l'information sont spécifiques des Eucaryotes. Les **petits ARN nucléaires (ARNpn ; en anglais ARNsn pour small nuclear RNA)** sont impliqués dans l'épissage des pré-ARNm en ARN messagers dans le noyau des cellules eucaryotes. Plusieurs ARNpn différents, associés à plusieurs sous-unités protéiques, forment un assemblage macromoléculaire appelé une PRNpn (petite particule ribonucléoprotéique nucléaire ; en anglais snRNP pour small nuclear ribonucleoprotein particle). Les **petits ARN cytoplasmiques (ARNpc)** sont impliqués dans le mouvement des protéines à l'intérieur de la cellule eucaryote. Plus spécifiquement, ils garantissent que des polypeptides destinés par exemple à être sécrétés hors de la cellule, sont insérés dans un des compartiments membranaires de la cellule (le reticulum endoplasmique rugueux). Ceci débute le processus de sécrétion protéique.

MESSAGE

Il existe deux types de gènes, ceux qui codent des protéines (la majorité) et ceux qui codent des ARN fonctionnels.

Toutes les fonctions des ADN et des ARN sont basées sur deux éléments clés :

1. Les bases complémentaires de chaînes nucléotidiques simple-brin peuvent s'associer par des liaisons hydrogène pour former des structures double-brin.
2. Des séquences de bases particulières appartenant à des acides nucléiques simple-brin ou double-brin peuvent être reconnues par des protéines spécifiques qui se lient aux acides nucléiques. (Ces protéines sont elles-mêmes codées par des gènes, mais leur fonction est d'agir sur d'autres gènes.)

Nous verrons ces deux propriétés en action dans les discussions détaillées qui suivent sur la transcription et la traduction.

MESSAGE

Les opérations sur l'ADN et sur l'ARN sont basées sur la complémentarité des séquences des bases et sur la fixation à des sites spécifiques, de diverses protéines de liaison aux acides nucléiques.

FABRIQUER DES TRANSCRITS FONCTIONNELS

Comment une séquence codant un gène, qui est un segment relativement court d'ADN englobé dans une molécule d'ADN bien plus longue (le chromosome), est-elle transcrite en une molécule d'ARN simple-brin de longueur et de séquence nucléotidique exactes ? Notre étude des différentes étapes de la transcription commencera par l'examen de la nature de la matrice.

La transcription utilise un brin d'ADN comme matrice

La transcription repose sur l'appariement des bases complémentaires. Les deux brins d'ADN de la double hélice se séparent localement et l'un des brins séparés sert de **matrice** (de guide) pour la synthèse d'ARN. Sur la totalité d'un chromosome, les deux brins d'ADN servent de matrice, *mais pour chaque gène, un seul brin est utilisé*, et il s'agit toujours du même brin pour ce gène (figure 3-2).

Ensuite, des nucléotides libres qui ont été synthétisés dans la cellule s'alignent sur la matrice d'ADN, « appelés » là en raison de leur complémentarité avec les bases de la matrice. Les ribonucléotides libres A s'alignent face aux T dans l'ADN, les G face aux C, les C face aux G et les U face aux A. Ce processus est catalysé par l'enzyme **ARN polymérase**, qui se fixe et se déplace le long de l'ADN en ajoutant des ribonucléotides à la chaîne d'ARN en cours d'élongation, comme le montre la figure 3-3a*. Nous voyons donc ici les deux principes de complémentarité des bases et de fixation des protéines (dans ce cas l'ARN polymérase) à l'œuvre.

L'ARN s'allonge toujours dans le sens 5'→3' ; en d'autres termes, les nucléotides sont toujours ajoutés à l'extrémité 3', comme le montre la figure 3-3b. En raison de la nature antiparallèle de l'appariement des nucléotides, le fait que l'ARN soit synthétisé dans le sens 5'→3' signifie que le brin matrice doit être orienté dans le sens 3'→5'. Lorsque les molécules d'ARN polymérase se déplacent le long du gène, on peut voir les molécules d'ARN en cours de synthèse au microscope électronique, comme dans la figure 3-4.

La complémentarité des bases signifie également que la séquence nucléotidique de l'ARN est la même que celle du brin qui ne sert pas de matrice, excepté le fait que les T sont remplacés par des U, comme le montre la figure 3-5 à la page 56. Le brin complémentaire du brin matrice correspond au transcrit d'ARN dans son orientation et dans sa séquence de bases (ce qui facilite les comparaisons). Lorsque l'on désigne des séquences d'ADN, c'est par

Ndlr : Un modèle alternatif a toutefois été récemment proposé. Voir pour revue l'article • The organization of replication and transcription • Peter R. Cook, *Science*, 1999, 284, pp 1790-1795.

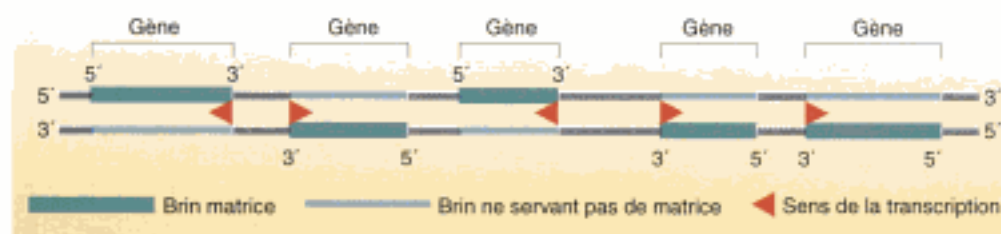
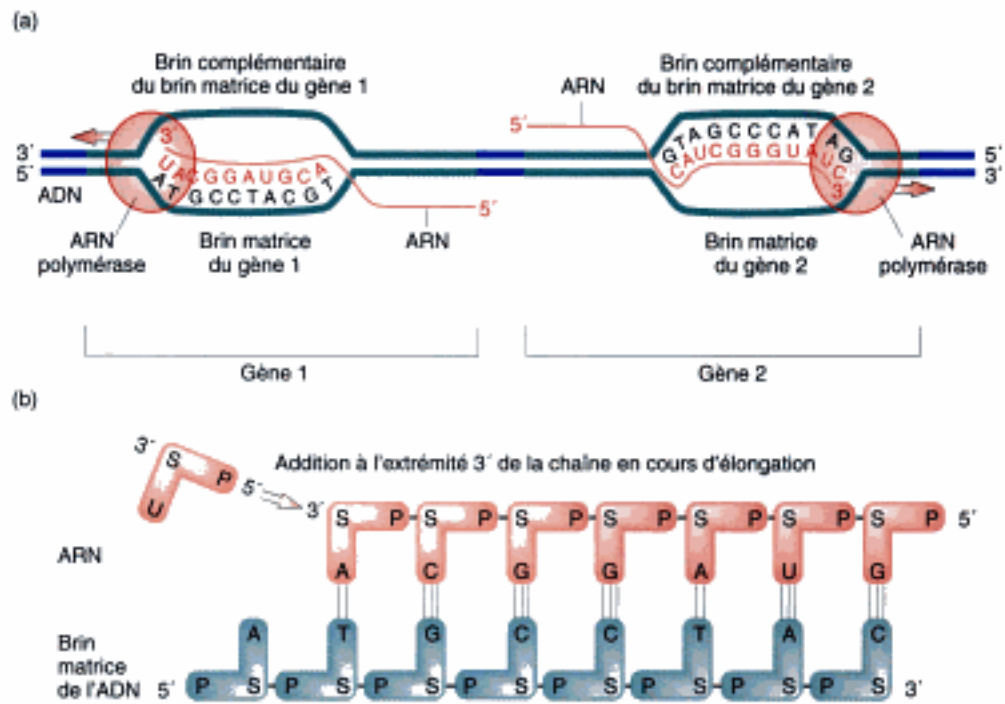


Figure 3-2 Brins d'ADN utilisés comme matrices pour la transcription.

Figure 3-3 Transcription de deux gènes. (a) L'ARN polymérase se déplace à partir de l'extrémité 3' du brin matrice, créant un brin d'ARN qui s'allonge dans le sens 5'→3' (car il doit être antiparallèle au brin matrice). Remarquez que certains gènes sont transcrits à partir d'un brin de la double hélice tandis que d'autres gènes utilisent l'autre brin comme matrice. (b) Un uracile est ajouté à l'extrémité 3' du transcrit pour le gène 1. La croissance se fait donc dans le sens 5'→3'.



convention la séquence du brin complémentaire du brin matrice qui est donnée, *bien que l'ARN soit transcrit à partir du brin matrice*. Cette distinction est extrêmement importante à garder à l'esprit au cours des discussions sur la transcription.

MESSAGE

La transcription est asymétrique : un seul brin de l'ADN d'un gène est utilisé comme matrice pour la transcription. Ce brin est celui dont l'orientation est 3'→5', et l'ARN est synthétisé dans le sens 5'→3'.

Les ARN polymérases

Chez la plupart des Procaryotes, c'est la même ARN polymérase qui transcrit tous les types d'ARN. Chez les Eucaryotes en revanche, il existe trois ARN polymérases différentes, qui sont spécialisées de la manière suivante :

1. L'ARN polymérase I (Pol I) transcrit les gènes d'ARNr.
2. L'ARN polymérase II (Pol II) transcrit les gènes codant des protéines.
3. L'ARN polymérase III (Pol III) transcrit les autres gènes d'ARN fonctionnels (par exemple les gènes d'ARNt).

Chez les Eucaryotes, la transcription des chromosomes du noyau se déroule intégralement dans ce compartiment. Les transcrits sortent ensuite du noyau par les pores nucléaires et gagnent le cytosol (la phase liquide du cytoplasme), où a lieu la traduction. Les génomes des

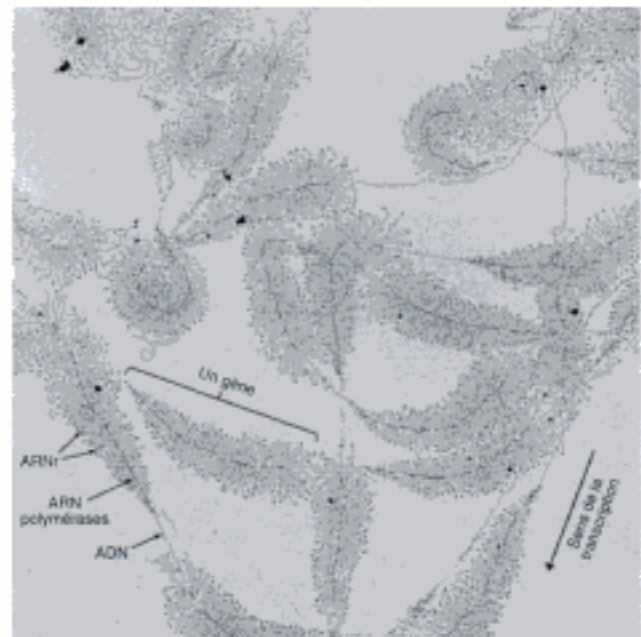


Figure 3-4 Gènes d'ARN ribosomiaux (ARNr) répétés en tandem en cours de transcription dans le nucléole de *Triturus viridescens* (un amphibien). (L'ARNr est un composant des ribosomes.) Le long de chaque gène, de nombreuses molécules d'ARN polymérase sont fixées et effectuent leur transcription dans le même sens. Les molécules d'ARN en cours d'élongation ressemblent à des fils qui s'allongent à partir du squelette d'ADN. Les molécules plus courtes d'ARN sont situées plus près du site d'initiation de la transcription ; les plus longues sont presque terminées. D'où l'aspect en « sapin de Noël ». (Photographie de O. L. Miller, Jr. et Barbara A. Hamkalo.)

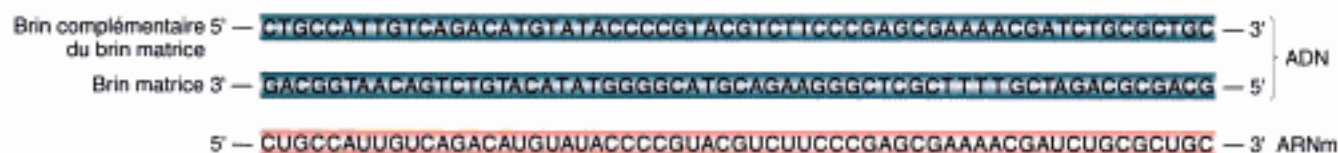


Figure 3-5 La séquence d'ARNm est complémentaire du brin matrice d'ADN à partir duquel il est synthétisé. La séquence présentée ici appartient au gène de l'enzyme β -galactosidase, qui est impliquée dans le métabolisme du lactose.

organites sont transcrits à l'intérieur de leur organe et la traduction s'y déroule également. Puisque les Procaryotes n'ont pas de noyau, il n'existe pas de déplacement comparable des transcrits et la traduction peut avoir lieu immédiatement, sur le transcrit en cours d'élongation.

À présent que nous avons introduit les acteurs, regardons les trois étapes de la transcription : l'initiation, l'élongation et la terminaison, en utilisant comme exemple, le système de la bactérie de l'intestin *E. coli*.

L'initiation

Une séquence d'ADN à laquelle se fixe une ARN polymérase pour initier la transcription est appelée un **promoteur**. Un promoteur fait partie de la région régulatrice adjacente à la région codante d'un gène. Rappelons-nous, puisqu'un transcrit d'ARN est synthétisé dans le sens 5'→3', on examine par convention le gène également dans le sens 5'→3' (l'orientation du brin complémentaire du brin matrice), même si la transcription commence effectivement à l'extrémité 3' du brin matrice. Par convention l'extrémité du gène transcrite en premier est appelée l'extrémité 5'. En d'autres termes, le gène est décrit de la même façon que son transcrit en ARN. En utilisant ce principe, le promoteur se trouve au début du gène. On dit donc qu'il est à l'extrémité 5' du gène et on appelle la région régulatrice, région régulatrice 5' (figure 3-6).

La figure 3-7 présente les séquences promotrices de 13 points d'initiation différents de la transcription dans le génome d'*E. coli*. Le schéma présente les séquences de bases qui apparaissent juste avant la première base transcrite, que l'on appelle le « site d'initiation ». Deux régions partiellement semblables apparaissent dans presque tous les cas. On a appelé ces régions, *régions* -35 (moins 35) et -10 en raison de leur position par rapport au point d'initiation de la transcription. Au bas de la figure 3-7 est indiquée la séquence consensus du promoteur : en d'autres termes, la séquence « moyenne » du promoteur. L'ARN polymérase parcourt l'ADN à la recherche d'une séquence promotrice, se fixe à l'ADN à cet endroit, le déroule et commence la synthèse d'une molécule d'ARN au niveau du site d'initiation de la transcription. Nous voyons donc que le principe de liaison à l'ADN résulte d'interactions entre la protéine (ici l'ARN polymérase) et une séquence spécifique de bases dans l'ADN.

L'élongation

L'ARN polymérase avance le long de l'ADN, maintenant ouverte une « bulle » de transcription pour exposer le brin matrice, et catalyse l'élongation en 3' du brin d'ARN en cours de synthèse. La polymérase compare les ribonucléotides triphosphate libres avec la base exposée qui suit dans l'ADN matrice et, lorsqu'il y a complémentarité,

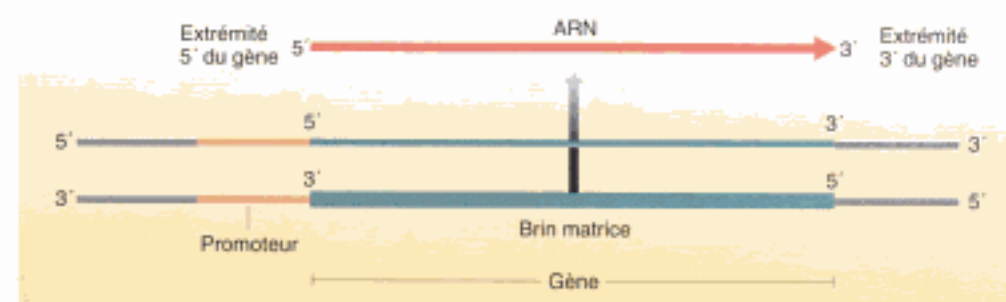


Figure 3-6 Convention utilisée pour désigner les extrémités 5' et 3' d'un gène.

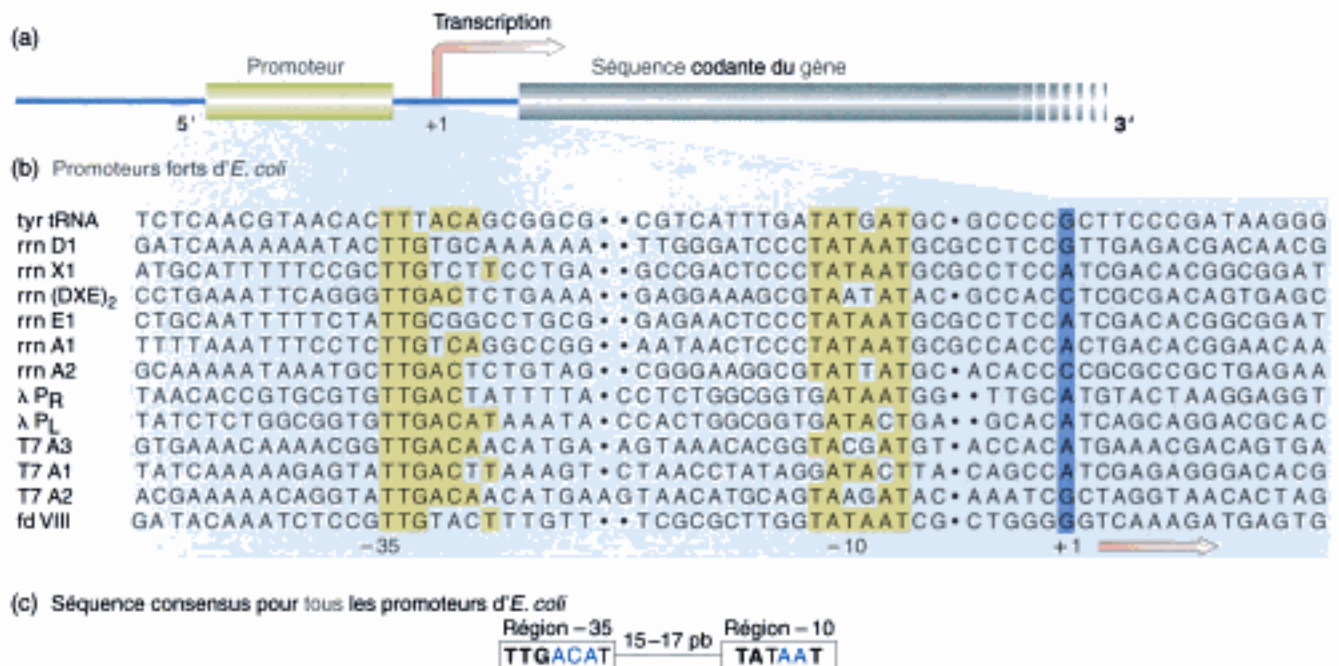


Figure 3-7 Séquence promotrice. (a) Le promoteur est situé « en amont » (vers l'extrémité 5') du site d'initiation et des séquences codantes. (Rappelez-vous que la transcription a lieu en fait sur le brin complémentaire). (b) Les sites promoteurs contiennent des régions de séquences similaires, comme l'indique la région en jaune de 13 séquences promotrices différentes chez *E. coli*. Les espaces (points) ont été ajoutés pour maximiser l'homologie des séquences consensus. Le gène gouverné par chaque séquence promotrice est indiqué sur la gauche. Les nombres correspondent au nombre de paires de bases avant (-) ou après (+) le site d'initiation de la synthèse d'ARN. (c) Le code des couleurs des séquences consensus de tous les promoteurs d'*E. coli* est le suivant : lettres bleues, > 75 % d'homologie ; lettres noires en gras, de 50 à 75 % ; lettres noires, de 40 à 50 %. [D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipurski, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997. Voir aussi W. R. McClure, *Ann. Rev. Biochem.* 54, 1985, 171 (Séquences consensus).]

ajoute le ribonucléotide correspondant à la chaîne. L'énergie de la réaction est fournie par la coupure du triphosphate riche en énergie, en monophosphate et en diphosphate inorganique selon la formule générale suivante



La terminaison

Lorsque l'ARN polymérase reconnaît des séquences nucléotidiques spécifiques dans l'ADN qui agissent comme des signaux de terminaison de la synthèse de la chaîne, le brin d'ARN et la polymérase sont libérés de la matrice d'ADN. Il existe deux mécanismes principaux de terminaison chez *E. coli*. Nous ne nous intéresserons ici qu'au mécanisme de terminaison direct. Les séquences de terminaison sont constituées d'environ 40 pb et se terminent par une séquence riche en GC suivie d'une succession de six

A, ou plus, sur le brin matrice. Les séquences GC correspondantes sur l'ARN sont arrangées de telle sorte que le transcrit est capable de former avec lui-même des liaisons complémentaires dans cette région, comme on peut le voir dans la figure 3-8. On appelle la région d'ARN

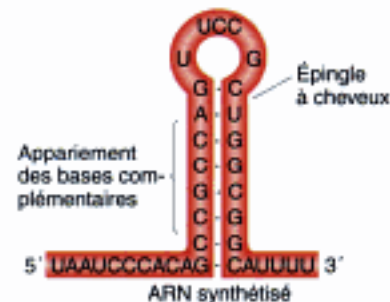


Figure 3-8 Structure du site de terminaison pour l'ARN polymérase bactérienne. La structure en épingle à cheveux se forme par appariement des bases complémentaires appartenant au même brin d'ARN.

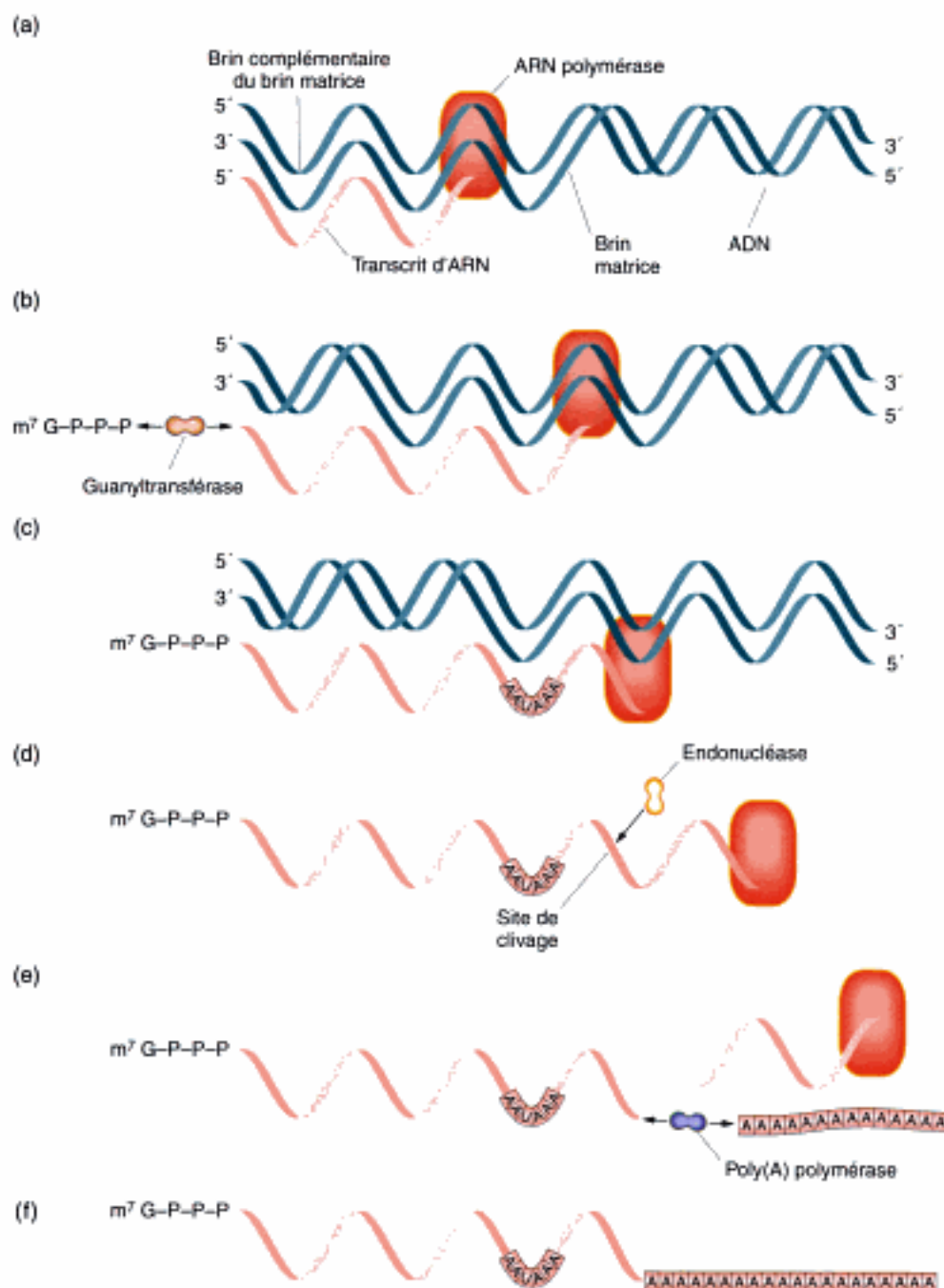


Figure 3-9 Maturation d'un transcrit primaire. (a) La transcription se fait par l'intermédiaire de l'ARN polymérase. (b) Au début de la transcription, une enzyme, la guanylyltransférase, ajoute une 7-méthylguanosine (m^7 Gppp) à l'extrémité 5' de l'ARNm. (c) La séquence AAUAAA, à proximité de l'extrémité 3' est le signal d'un événement de clivage (d) par une endonucléase, qui se produit approximativement à 20 pb en aval. (e) Une enzyme, la poly(A) polymérase ajoute ensuite une queue poly(A), constituée de 150 à 200 résidus d'adénosine, au site de ce clivage, à l'extrémité 3'. (f) Ceci produit l'ARN primaire complet. (D'après J. E. Darnell, Jr., • The processing of RNA • Copyright © 1983 par Scientific American, Inc. Tous droits réservés.)

double-brin résultante, une **épingle à cheveux**. Elle se poursuit et se termine par une succession de U qui correspond aux résidus A sur la matrice d'ADN. L'épingle à cheveux et la succession de résidus U semblent servir de signal pour la libération de l'ARN polymérase et la termi-

cription, une **coiffe** (cap) constituée d'un résidu 7-méthyl-guanosine est ajoutée à l'extrémité 5' du transcrit, par l'intermédiaire d'une liaison triphosphate. Puis une séquence AAUAAA située à proximité de l'extrémité 3' du transcrit sert de signal pour la terminaison de la transcription par une enzyme qui coupe la chaîne d'ARN.

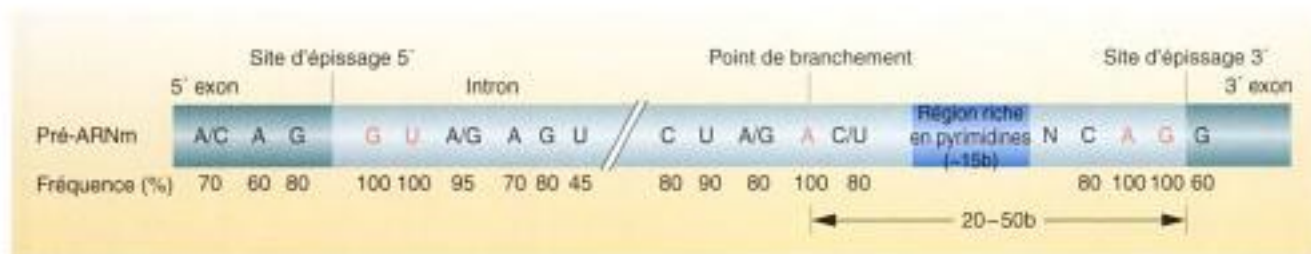


Figure 3-11 Séquences conservées en rapport avec l'épissage des introns. Les nombres indiqués sous les nucléotides représentent le pourcentage de conservation.

dont les morceaux peuvent être coupés et recollés. L'épissage enlève les introns et joint les régions codantes, les exons, pour former un ARNm, qui est à présent constitué d'une séquence parfaitement colinéaire avec celle de la protéine.

Le mécanisme d'épissage des introns. Les jonctions exon-intron des ARNm présentent des séquences spécifiques qui sont hautement conservées. En d'autres termes, elles sont identiques pour la majorité des introns chez la plupart des espèces. Ces séquences jouent un rôle dans les réactions d'épissage. Par exemple, il y a un GU au site d'épissage 5' de l'intron et un AG au site d'épissage 3' dans quasiment tous les cas examinés (• **la règle GU-AG** • ; voir figure 3-11). On trouve d'autres séquences conservées de façon moins stricte de part et d'autre de celles-ci. Ces configurations courantes du pré-ARNm sont reconnues par de petites particules ribonucléoprotéiques ou RNPPn qui catalysent les réactions d'excision et d'épissage. Les RNPPn sont des complexes de protéines et de petits ARN nucléaires (ARNpn). Grâce à la complémentarité des bases de leurs ARNpn, les RNPPn peuvent, au moyen de liaisons hydrogène, aider à aligner les sites d'épissage avec les séquences situées aux limites exon-intron (figure 3-12). L'intron est ensuite excisé. Un intermédiaire en forme de lasso est formé puis libéré pour produire l'ARNm final. La figure 3-13 représente le modèle en lasso. Ici le site d'épissage 5' est d'abord clivé [figure 3-13(a)] au cours d'une réaction qui crée une boucle [figure 3-13(b)]. Le groupement 3'-OH qui vient d'être formé clive ensuite le site d'épissage 3' [figure 3-13(b)], libérant l'intron partiellement enroulé et reliant au même moment les deux exons [figure 3-13(c)]. La formation du lasso exige elle aussi une séquence conservée dans l'intron et la fixation a toujours lieu au niveau d'une adénine hautement conservée.

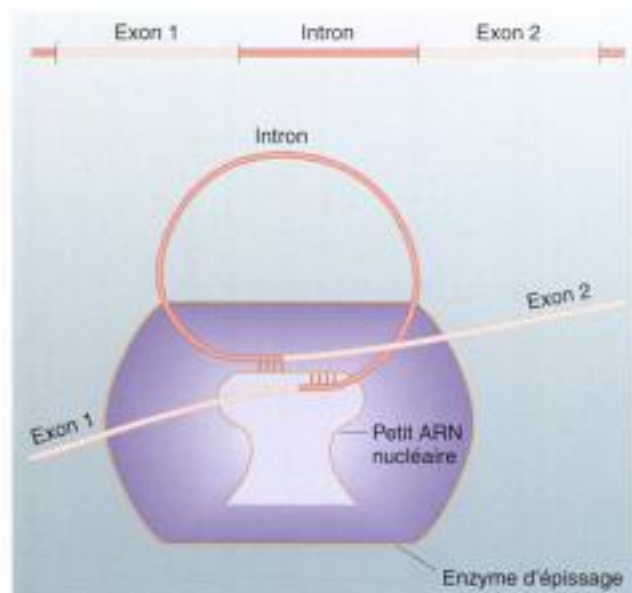


Figure 3-12 Schéma décrivant la participation possible de petits ARN nucléaires dans la réaction d'épissage.

Au cours du processus d'épissage, les RNPPn, le transcrit primaire et les facteurs associés, se combinent pour former un complexe ribonucléoprotéique de masse moléculaire élevée (60S) appelé un **spliceosome** (figure 3-14) qui catalyse les réactions d'épissage décrites ci-dessus.

Un cas exceptionnel d'épissage de l'ARN se produit chez le protozoaire cilié *Tetrahymena* chez lequel la réaction d'épissage est catalysée par la molécule d'ARN transcrite, elle-même. Cette découverte extraordinaire a permis de démontrer pour la première fois qu'une molécule d'ARN peut dans certains cas agir comme une enzyme et catalyser une réaction biologique spécifique. Ces ARN dotés d'une activité enzymatique ont été appelés **ribozymes**.

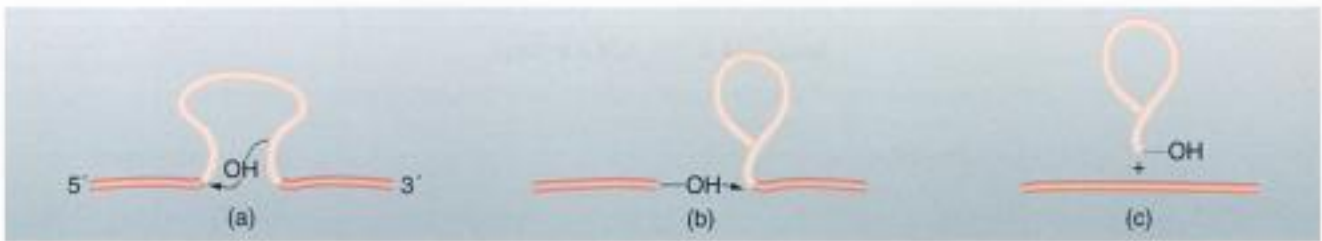


Figure 3-13 Lasso ou boucle formé(e) par une catégorie d'introns lorsqu'ils sont excisés de leur molécule d'ARN. (a) L'un des nombreux groupements 2'-hydroxyle de l'intron attaque le site d'épissage 5'. (b) La réaction suivante relie donc l'extrémité 5' de l'intron non pas à l'extrémité 3' mais au point de l'attaque du groupement 2'-hydroxyle un peu plus loin, formant une structure avec un branchement comportant une boucle : le lasso. (c) Le branchement se produit par formation d'une nouvelle liaison phosphodiester 2'-5' qui permet à un nucléotide adénosine de former des liaisons phosphodiester avec trois autres nucléotides plutôt qu'avec les deux habituels. La réunion des exons libère le lasso du reste de l'ARN. (D'après T. Cech, « RNA as an enzyme ». Copyright © 1986 par Scientific American, Inc. Tous droits réservés.)

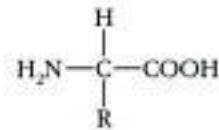
MESSAGE

L'ARNm eucaryote subit une maturation par addition d'une coiffe en 5', par polyadénylation en 3' et par excision de ses introns.

Une fois qu'un transcrit primaire a subi une maturation complète en une molécule d'ARNm mature, la traduction en protéine peut avoir lieu. Avant d'en venir au mécanisme de la traduction, nous devons d'abord comprendre la nature de la structure d'une protéine.

LA STRUCTURE D'UNE PROTÉINE

Une protéine est un polymère constitué de monomères appelés **acides aminés**. En d'autres termes, il s'agit d'une chaîne d'acides aminés que l'on désigne parfois par le terme de **polypeptide**. Les acides aminés ont la formule générale suivante :



La chaîne latérale, ou groupement R (réactif) est très variable, d'un atome d'hydrogène (comme pour l'acide aminé appelé glycine) à un cycle complexe (comme pour le tryptophane, un autre acide aminé). On dénombre 20 acides aminés entrant dans la composition des protéines (voir tableau 3-1 à la page suivante), chacun ayant un groupement R différent, ce qui donne à l'acide aminé ses propriétés uniques. Dans les protéines, les acides aminés sont liés les uns aux autres par des liaisons covalentes appelées **liaisons peptidiques**. Une liaison peptidique se forme par une réaction de condensation au cours de

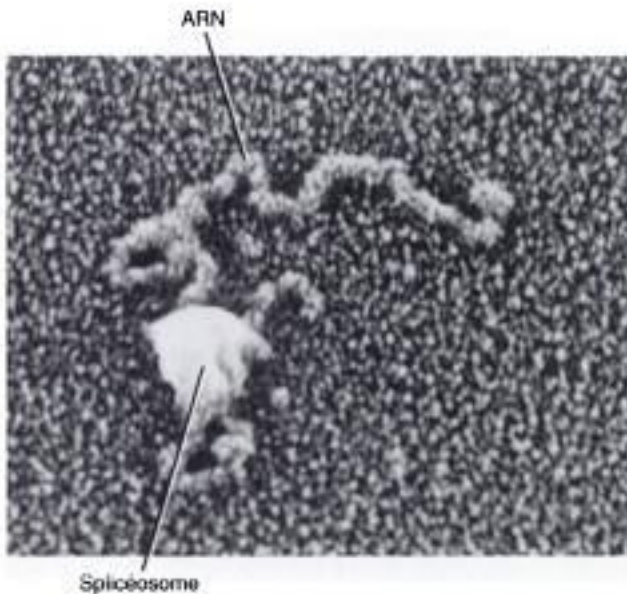
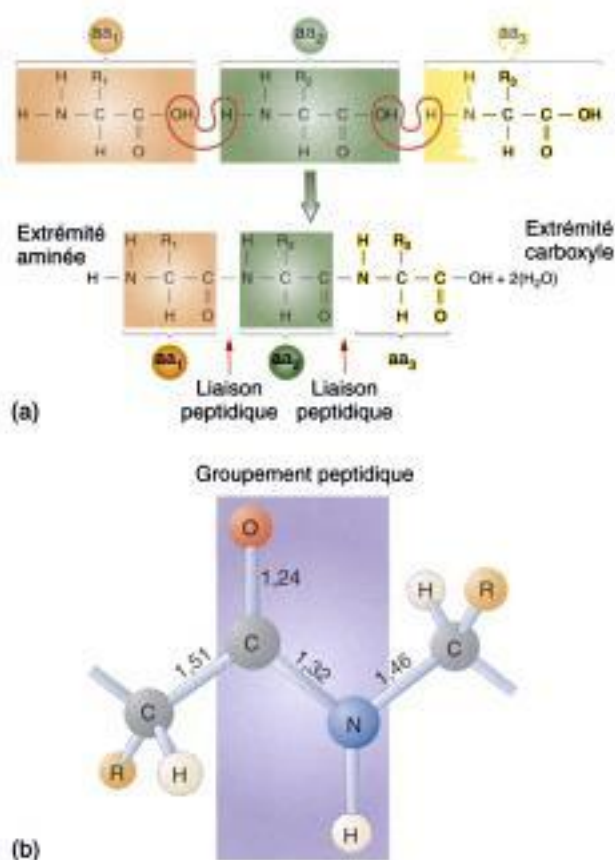


Figure 3-14 Micrographie électronique d'un splicéosome. (D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipurski, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997.)

TABLEAU 3-1

Les 20 acides aminés habituels chez les organismes vivants

Abréviation			Abréviation		
Acide aminé	3 lettres	1 lettre	Acide aminé	3 lettres	1 lettre
Acide aspartique	Asp	D	Leucine	Leu	L
Acide glutamique	Glu	E	Lysine	Lys	K
Alanine	Ala	A	Méthionine	Met	M
Arginine	Arg	R	Phénylalanine	Phe	F
Asparagine	Asn	N	Proline	Pro	P
Cystéine	Cys	C	Sérine	Ser	S
Glutamine	Gln	Q	Thréonine	Thr	T
Glycine	Gly	G	Tryptophane	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

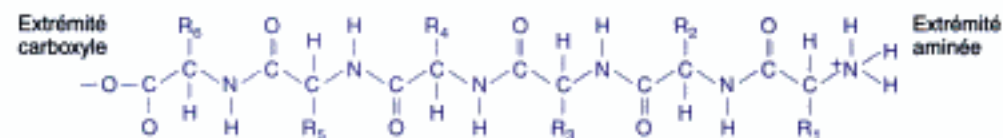


laquelle une molécule d'eau est enlevée (figure 3-15). En raison de la façon dont se forme la liaison peptidique, une chaîne polypeptidique possède toujours une **extrémité aminée** (NH₂) et une **extrémité carboxyle** (COOH), comme on le voit dans cette figure.

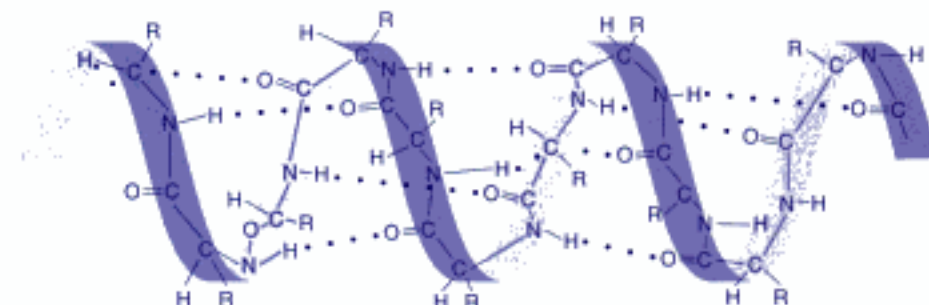
Les protéines ont une structure complexe à quatre niveaux d'organisation (figure 3-16). La séquence linéaire d'acides aminés dans une chaîne polypeptidique constitue la **structure primaire** de la protéine. Plusieurs sortes de liaisons faibles (liaisons hydrogène, électrostatiques, et de van der Waals) établies entre les atomes de la chaîne polypeptidique ou entre des chaînes polypeptidiques différentes, confèrent à la protéine une forme spécifique. La **structure secondaire** d'une protéine apparaît à la suite des interactions des acides aminés qui sont proches les uns des autres dans la séquence linéaire, formant le plus

Figure 3-15 La liaison peptidique. (a) Un polypeptide est formé par le retrait d'une molécule d'eau entre des acides aminés pour former des liaisons peptidiques. Chaque « aa » indique un acide aminé. R₁, R₂ et R₃ représentent des groupes R (chaînes latérales) qui différencient les acides aminés les uns des autres. R peut tout représenter, d'un atome d'hydrogène (comme dans la glycine) à un cycle complexe (comme dans le tryptophane). (b) Le groupement peptidique est une unité plane rigide avec des groupements R qui saillent à partir du squelette C-N. Les distances standard des liaisons (en angströms) sont indiquées. (Partie b d'après L. Stryer, *Biochemistry*, 4^e éd. Copyright © 1995 W. H. Freeman and Company.)

(a) Structure primaire



(b) Structure secondaire



Liaisons hydrogène entre des acides aminés situés en différentes positions de la chaîne polypeptidique

Figure 3-16 Différents niveaux de structure protéique.

(a) Structure primaire.

(b) Structure secondaire.

Le polypeptide présenté en partie (a) est dessiné sous la forme d'une hélice α à la suite de la formation de liaisons hydrogène.

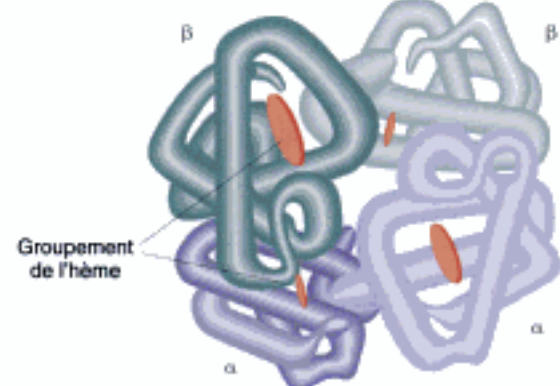
(c) Structure tertiaire : la structure tridimensionnelle de la myoglobine.

(d) Structure quaternaire : l'arrangement de deux sous-unités α et de deux sous-unités β pour former la structure quaternaire complète d'une autre protéine, l'hémoglobine.

(c) Structure tertiaire



(d) Structure quaternaire



souvent une hélice α . La **structure tertiaire** est produite par le repliement de l'hélice, ou d'une structure secondaire d'un autre type. Dans certaines protéines, deux structures tertiaires ou plus se rassemblent en une **structure quaternaire**. L'association quaternaire peut avoir lieu entre différents types de polypeptides ou entre des polypeptides identiques.

De nombreuses protéines ont des structures compactes et sont appelées **protéines globulaires**. Les enzymes et les anticorps sont parmi les protéines globulaires les plus importantes. Les protéines présentant une forme linéaire, appelées **protéines fibreuses**, sont des composants importants de structures telles que les cheveux et les muscles.

La forme est de toute première importance pour une protéine car c'est cette forme particulière qui lui permet

d'accomplir sa tâche spécifique dans la cellule. Par exemple, une enzyme a une poche spécifique (le **site actif**) dans laquelle son substrat peut venir se loger (voir figure 3-17 à la page suivante). La forme est en grande partie déterminée par la séquence primaire d'acides aminés. En dernier lieu, même la structure quaternaire d'une protéine est déterminée par les structures primaires des chaînes d'acides aminés qui la composent. Toutefois, on ne comprend pas encore parfaitement les règles selon lesquelles la structure primaire est convertie en structures d'ordre supérieur. La séquence d'acides aminés détermine également les groupements R disponibles pour permettre à la protéine de se fixer à d'autres composants cellulaires spécifiques.



Figure 3-17 Représentation de la structure de la transcriptase inverse, montrant son site actif, dans lequel l'ARN est fixé et copié en ADN. (D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipurski, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997.)

LA TRADUCTION

Comment une séquence d'acides aminés est-elle déterminée ? Très simplement, par la séquence nucléotidique. La séquence d'acides aminés spécifique d'un polypeptide est déterminée par la séquence nucléotidique du gène qui la code. La séquence de nucléotides de l'ADN d'un gène est transcrite en une séquence équivalente d'ARNm. Rappelons-nous que c'est le brin matrice d'ADN qui est transcrit mais que l'ARNm synthétisé a la même séquence que l'autre brin, le brin complémentaire. Un groupe de ribosomes parcourt l'ARNm, chacun débutant à l'extrémité 5' et avançant sur toute la longueur de l'ARNm pour gagner l'extrémité 3'. Au fur et à mesure qu'un ribosome avance, il « lit » la séquence nucléotidique de l'ARNm trois nucléotides à la fois. Chaque groupe de trois est appelé un triplet ou **codon** et correspond à un acide aminé spécifique. Puisqu'il y a 4 nucléotides différents dans l'ARNm, il y a $4 \times 4 \times 4 = 64$ codons différents possibles. Ces codons et les acides aminés auxquels ils correspondent sont présentés dans la figure 3-18. Le même **code génétique** est utilisé par quasiment tous les organismes de la planète. Il existe des exceptions pour lesquelles quelques codons ont une signification différente.

	Deuxième lettre				Troisième lettre		
	U	C	A		G	U	C
Première lettre	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U	C	A
	UUC	UCC	UAC	UGC	U	C	A
	UUA Leu	UCA	UAA Stop	UGA Stop	U	C	A
	UUG	UCG	UAG Stop	UGG Trp	U	C	A
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U	C	A
	CUC	CCC	CAC	CGC	U	C	A
	CUA	CCA	CAA Gln	CGA	U	C	A
	CUG	CCG	CAG	CGG	U	C	A
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U	C	A
	AUC	ACC	AAC	AGC	U	C	A
	AUA	ACA	AAA Lys	AGA Arg	U	C	A
	AUG Met	ACG	AAG	AGG	U	C	A
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U	C	A
	GUC	GCC	GAC	GGC	U	C	A
	GUA	GCA	GAA Glu	GGA	U	C	A
	GUG	GCG	GAG	GGG	U	C	A

Figure 3-18 Le code génétique.

Un codon est traduit en un acide aminé par un ARNt comportant une séquence complémentaire

La base de la spécificité entre codon et acide aminé se trouve dans la structure des molécules d'ARN de transfert (ARNt). Une molécule d'ARNt a la forme d'une feuille de trèfle composée de quatre tiges en double hélice et de trois boucles simple-brin (figure 3-19). La boucle du milieu porte un triplet de nucléotides appelé l'**anticodon**, dont la tâche est de se lier à un codon spécifique de l'ARNm grâce à un appariement spécifique ARN-ARN. Puisque les codons de l'ARNm sont lus dans le sens 5'→3', les anticodons sont orientés dans le sens 3'→5', comme le montre la figure 3-19. Chaque ARNt est spécifique d'un acide aminé et le transporte fixé à son extrémité libre 3'. Les acides aminés sont ajoutés à l'ARNt par des enzymes appelées **aminoacyl-ARNt synthétases**. A chaque acide aminé correspond une synthétase spécifique qui l'attache seulement aux ARNt qui reconnaissent les codons de cet acide aminé. Une fois encore, nous voyons le principe de complémentarité des acides nucléiques à l'œuvre, cette fois dans la liaison du codon (ARN) à l'anticodon (également de l'ARN).

La figure 3-18 montre que le nombre de codons pour un même acide aminé varie, allant de un (tryptophane, UGG) à six (sérine, UCU, UCC, UCA, UCG, AGU ou AGC). On ne sait pas exactement pourquoi le code génétique comporte une telle variation mais deux faits permettent de s'en faire une idée :

1. Certains acides aminés peuvent être apportés au ribosome par plusieurs types d'ARNt portant des anticodons différents, tandis que d'autres acides aminés sont apportés au ribosome par un seul ARNt.
2. Certaines espèces d'ARNt peuvent apporter leur acide aminé spécifique en réponse à plusieurs

codons, et non à un seul, grâce à un appariement lâche à l'extrémité 3' du codon et à l'extrémité 5' de l'anticodon. Cet appariement lâche est appelé le **flottement** (*wobble* en anglais).

Le flottement correspond à la situation dans laquelle le troisième nucléotide d'un anticodon (à l'extrémité 5') peut former deux associations (figure 3-20). Le nucléotide peut former des liaisons hydrogène non seulement avec son nucléotide complémentaire normal en troisième position du codon, mais également avec un nucléotide différent à la même position. Il existe des « règles de flottement » qui décident des nucléotides qui peuvent ou non former des associations par le biais de liaisons hydrogène,

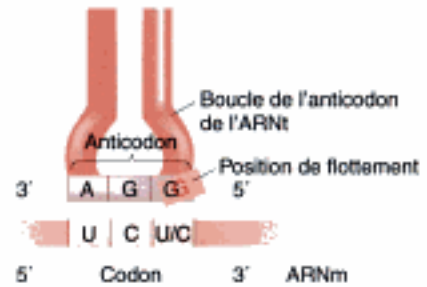


Figure 3-20 Au niveau du troisième site (extrémité 5') de l'anticodon, G peut adopter l'une ou l'autre des positions flottantes, et est donc capable de s'apparier soit avec U soit avec C. Ceci signifie qu'une seule espèce d'ARNt portant un acide aminé (dans ce cas de la sérine) peut reconnaître deux codons — UCU et UCC — de l'ARNm.

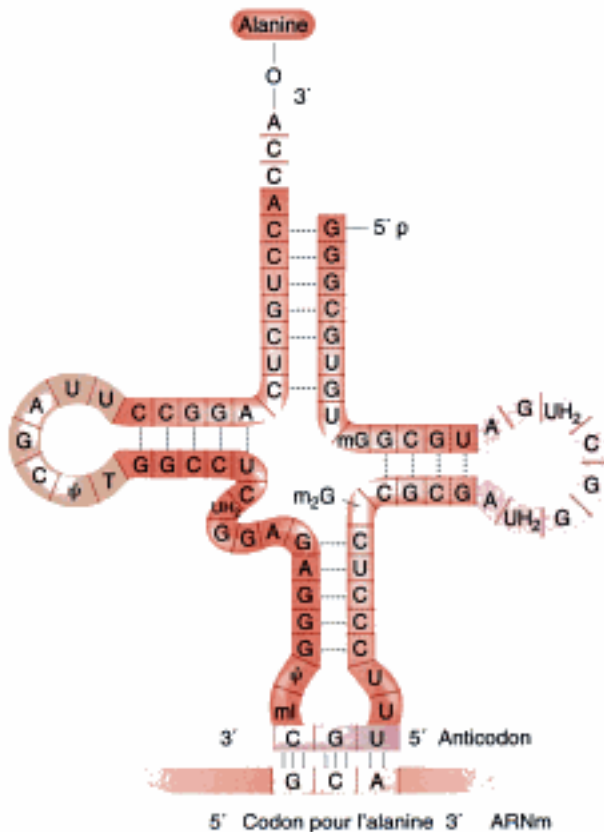


Figure 3-19 Structure de l'ARNt de l'alanine, montrant la liaison de l'aminocyl-ARNt avec le codon approprié de l'ARNm. Les flèches indiquent plusieurs types de bases rares modifiées. Les symboles ψ , mG, m₂G, ml et DHU (ou UH₂) sont des abréviations des bases modifiées : pseudouridine, méthylguanosine, diméthylguanosine, méthylinosine et dihydrouridine respectivement. (Adapté de L. Stryer, *Biochemistry*, 4^e éd. Copyright © 1995, par W. H. Freeman et Company.)

grâce au flottement (tableau 3-2). Dans le tableau, la lettre *I* désigne l'inosine, l'une des bases rares que l'on trouve dans l'ARNt, souvent dans l'anticodon.

Le tableau 3-3, à la page suivante, dresse la liste de tous les codons spécifiant la sérine et montre de quelle façon trois ARNt différents (ARNt^{Ser}₁, ARNt^{Ser}₂, ARNt^{Ser}₃) peuvent s'apparier avec ces codons. Dans certaines espèces, il existe une autre sorte d'ARNt que l'on pourrait représenter par ARNt^{Ser}₄, qui possède un anticodon identique à l'un des trois anticodons indiqués dans le tableau 3-3 mais dont la séquence nucléotidique diffère ailleurs. On appelle ces quatre ARNt des **ARNt isoaccepteurs** car ils acceptent le même acide aminé, mais sont transcrits à partir de gènes d'ARNt différents.

TABEAU 3-2

Appariements codon-anticodon permis par les règles de flottement, c'est-à-dire la flexibilité de la fixation.

Extrémité 5' de l'anticodon	Extrémité 3' du codon
G	U ou C
C	C seulement
A	U seulement
U	A ou G
I	U, C, ou A

TABLEAU 3-3

Les différents ARNt capables de reconnaître les codons de la sérine

ARNt	Anticodon	Codon
ARNt ^{Ser} ₁	AGG + flottement	UCU UCC
ARNt ^{Ser} ₂	AGU + flottement	UCA UCG
ARNt ^{Ser} ₃	UCG + flottement	AGU AGC

Les ribosomes sont les usines de fabrication des protéines

Le lieu de rencontre entre les acides aminés liés à leur ARNt et l'ARNm est le ribosome. Les ribosomes sont de gros assemblages macromoléculaires qui agissent comme des machines subcellulaires compliquées. Chaque ribosome est constitué d'une petite et d'une grande sous-unités. Chaque sous-unité est constituée elle-même de plusieurs types d'ARNr et jusqu'à 50 protéines (figure 3-21). Les ribosomes contiennent des sites spécifiques qui leur permettent de fixer l'ARNm, les ARNt et d'autres facteurs protéiques spécifiques, tous nécessaires à la synthèse protéique. Considérons d'abord la synthèse protéique sur le ribosome dans son ensemble, et examinons ensuite chacune des étapes du processus plus en détail.

La figure 3-22 montre un polypeptide en cours de

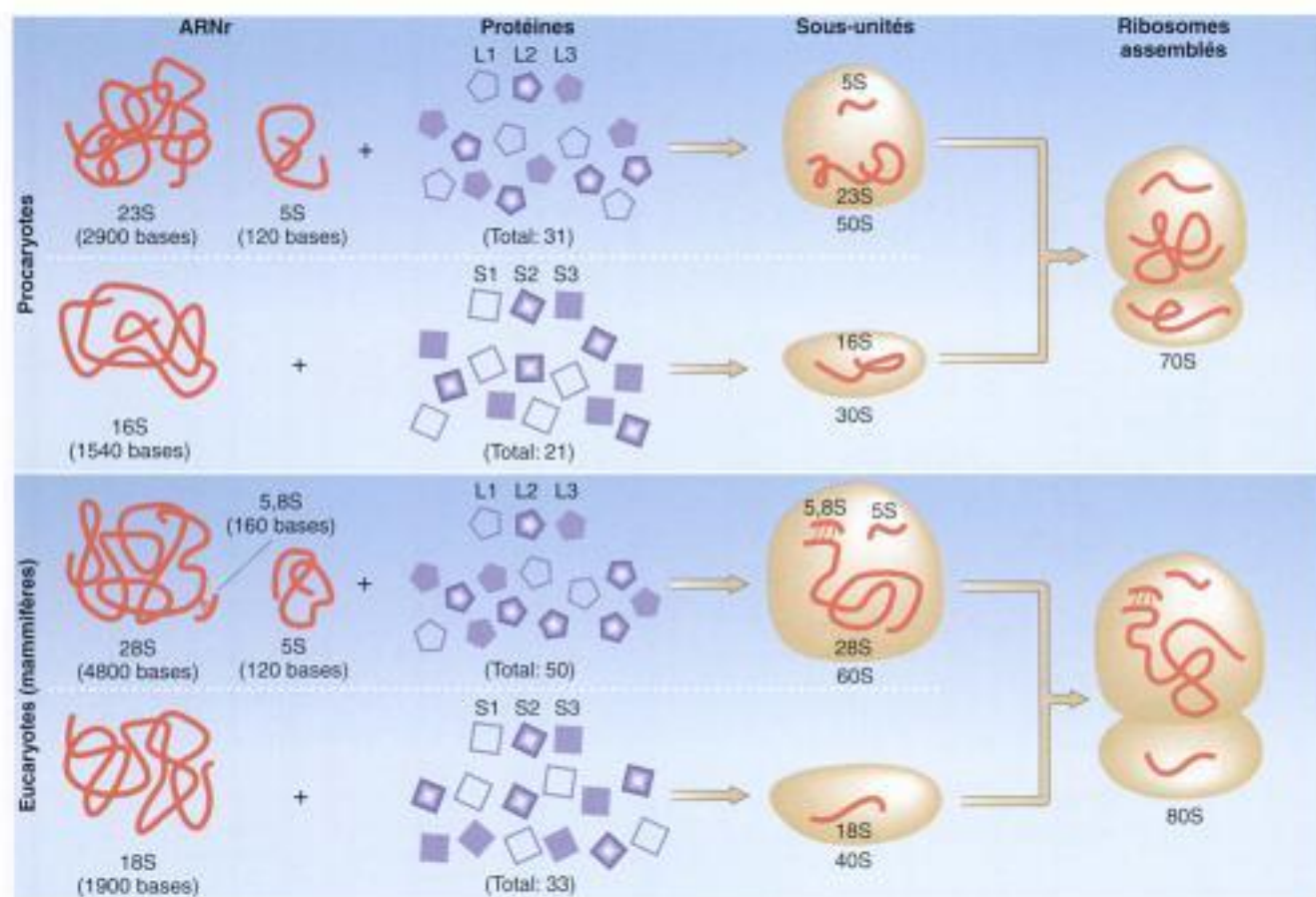


Figure 3-21 Les ribosomes contiennent une grande et une petite sous-unités. Chaque sous-unité comporte à la fois des ARNr de différentes longueurs et un ensemble de protéines (représentées par des couleurs différentes). Il y a deux molécules principales d'ARNr dans tous les ribosomes. Les ribosomes des Procaryotes contiennent également un ARNr de 120 bases de long qui sédimmente à 5S, tandis que les ribosomes eucaryotes comportent deux petits ARNr : une molécule d'ARNr 5S similaire à la 5S procaryote et une molécule de 5,8S de 160 bases de long. Les protéines de la grande (en anglais : *large*) sous-unité sont appelées L1, L2, etc et celles de la petite (*small*) sous-unité S1, S2, etc. (D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipurski, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997.)

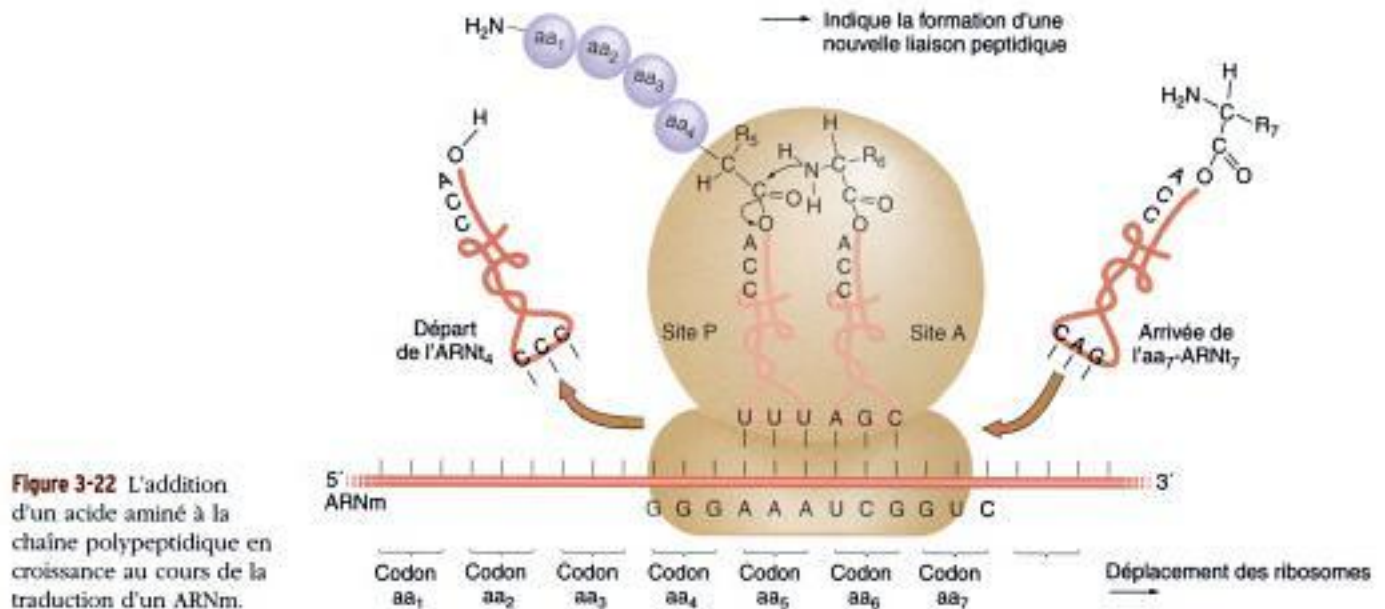


Figure 3-22 L'addition d'un acide aminé à la chaîne polypeptidique en croissance au cours de la traduction d'un ARNm.

synthèse sur le ribosome. L'ARNm se fixe à la petite sous-unité. Les ARNt se fixent sur deux sites qui chevauchent les sous-unités. Le **site A** est le site d'entrée pour un aminoacyl-ARNt (un ARNt portant un seul acide aminé). Le peptidyl-ARNt portant la chaîne polypeptidique en croissance se fixe au **site P**. Chaque nouvel acide aminé est ajouté par transfert de la chaîne en cours d'élongation au nouvel aminoacyl-ARNt, formant une nouvelle liaison peptidique. L'ARNt qui s'est séparé de son acide aminé est alors libéré du site P et le ribosome glisse un codon plus loin le long du messager, transférant le nouveau peptidyl-ARNt au site P et laissant le site A vacant pour l'acidoacyl-ARNt entrant suivant. L'extrémité croissante de la chaîne peptidique (c'est-à-dire l'extrémité sur le ribosome) est l'extrémité carboxyle et l'extrémité libre est l'extrémité aminée. L'extrémité aminée de la protéine correspond donc à l'extrémité 5' de l'ARNm et l'extrémité carboxyle à l'extrémité 3'.

L'action du ribosome au cours de la traduction est divisée en trois étapes distinctes : l'initiation, l'élongation et la terminaison.

L'initiation. Chez les bactéries, le premier acide aminé de tout peptide néo-synthétisé est la *N*-formylméthionine. Il est inséré non pas par l'ARNt^{Met}, mais par un ARNt initiateur appelé ARNt^{Met}. Cet ARNt initiateur a l'anticodon normal pour la méthionine mais insère une *N*-formylméthionine plutôt qu'une méthionine. Chez *E. coli*, AUG, GUG, et à de rares occasions UUG, servent de codons d'initiation. Lorsque l'un de ces codons se trouve à la position d'initiation, il est reconnu par un *N*-formylMet-ARNt et une *N*-formylméthionine apparaît alors comme premier acide aminé de la chaîne. (Le groupement formyle sera perdu plus tard.) Chez les Eucaryotes, AUG est généralement le

premier codon d'un polypeptide, mais c'est une méthionine qui est insérée et non une *N*-formylméthionine.

Comment les codons d'initiation corrects sont-ils sélectionnés au milieu de tous les codons AUG et GUG qui se trouvent dans une molécule d'ARNm ? Chez les bactéries, les codons d'initiation sont précédés de séquences appelées **séquences de Shine-Dalgarno** qui s'apparient avec l'extrémité 3' de l'ARNr 16S du ribosome, positionnant ainsi correctement le ribosome à côté du codon d'initiation. Il y a une région de séparation courte et variable entre la séquence de Shine-Dalgarno et le codon d'initiation.

Il est à noter que chez les Procaryotes, un ARNm comporte une « **séquence de tête** » non traduite (appelée aussi séquence leader : *leader sequence* en anglais) correspondant à l'ARN situé entre le site de début de la transcription et le site de début de la traduction (voir figure 3-7). Les ARNm eucaryotes possèdent également une séquence de tête mais il ne s'agit pas de la séquence de Shine-Dalgarno. Au lieu de cela, un ribosome portant un ARNt initiateur se fixe, guidé par la coiffe en 5', se déplace le long de l'ARN et initie la traduction au niveau du premier codon AUG adéquat qu'il rencontre. Chez les Eucaryotes, ce codon est déterminé par certains nucléotides conservés qui signalent si un AUG donné est le codon d'initiation de la traduction. Mais on ne sait pas si ces séquences sont reconnues par des interactions avec les ARN ribosomiaux.

L'élongation. Plusieurs protéines appelées *facteurs d'élongation*, qui guident la fixation et le déplacement des ARNt et du ribosome participent à l'élongation. L'énergie nécessaire à ces réactions est fournie par l'hydrolyse de GTP.

La terminaison. Vous avez peut-être remarqué dans la figure 3-18 que certains codons ne spécifient aucun acide aminé. Ces codons, UAG, UGA et UAA sont appelés **codons stop** ou **codons de terminaison**. On peut les considérer comme les signes de ponctuation qui terminent le message codé dans l'ARNm. Les codons stop sont souvent appelés *codons non-sens* car ils ne désignent aucun acide aminé. C'est trompeur, car la ponctuation constitue une partie importante du sens de tout message. Il est intéressant de constater que les trois codons stop ne sont pas reconnus par un ARNt mais par des facteurs protéiques appelés **facteurs de libération** (ou de relargage). Lorsque le peptidyl-ARNt se trouve dans le site P, les facteurs de libération se fixent au site A en réponse aux codons terminant la chaîne. Le polypeptide est alors libéré du site P et les ribosomes se dissocient en deux sous-unités. La traduction est alors achevée.

Une certaine façon d'analyser les séquences d'ADN par ordinateur à la recherche de gènes potentiels consiste à rechercher les **cadres de lecture ouverts** (ORF pour *open reading frame* en anglais), qui sont de longues séquences d'ADN commençant par un codon d'initiation (par exemple 5'-ATG-3') et se terminant par l'un des trois codons stop.

MESSAGE

La traduction est effectuée par des ribosomes qui se déplacent le long de l'ARNm dans le sens 5' → 3'. Un ensemble de molécules d'ARNt apporte les acides aminés jusqu'au ribosome en mouvement, leurs anticodons se fixant aux codons accessibles sur le ribosome.

FONCTIONNEMENT ET DYSFONCTIONNEMENT DES PROTÉINES DANS LES CELLULES

La transcription et la traduction garantissent qu'une succession linéaire de nucléotides de l'ADN d'un gène sera convertie en une succession linéaire d'acides aminés dans la chaîne polypeptidique primaire. Il existe une correspondance précise entre les codons de l'ADN et les acides aminés de la protéine. Cette correspondance linéaire entre le gène et le produit protéique est appelée **colinéarité**. La fonction de la protéine dépend donc intégralement de la séquence d'ADN du gène qui la code.

Les protéines peuvent être classées en deux grands types, les *protéines en tant qu'éléments actifs primordiaux* et les *protéines de structure*. Les **enzymes** sont de bons exemples de protéines « actives ». Ce sont les catalyseurs biologiques qui permettent les milliers de réactions chimiques qui se déroulent à l'intérieur d'une cellule vivante. Les gènes codant des enzymes exercent un important contrôle sur ce qui se passe à l'intérieur d'une cellule. Les protéines des microtubules, responsables des mouvements dans la cellule, et les protéines liées aux membranes, qui régulent le pompage des ions et des nutriments à travers

les membranes sont d'autres exemples de protéines « actives ». Les protéines structurales, comme leur nom l'indique, contribuent à l'architecture de la cellule et de l'organisme. La kératine contenue dans les cheveux et le collagène présent dans les os sont deux protéines structurales. Certaines protéines sont à la fois « actives » et structurales. Par exemple, l'actine et la myosine, qui sont des protéines contractiles, constituent une part importante des muscles et en affectent donc la forme. La fonction des protéines est bien illustrée par les enzymes.

La fonction des enzymes

Les sites actifs. Les enzymes sont des catalyseurs biologiques accélérant des réactions chimiques qui, sans cela, seraient trop lentes pour permettre à la cellule de fonctionner. Les molécules de substrat s'adaptent à la forme du site actif des enzymes. Les enzymes remplissent leur rôle en interagissant physiquement avec les molécules de substrat, pour créer ou rompre des liaisons chimiques. La figure 3-23 schématise le concept général et présente deux des modèles de l'action des enzymes. Les enzymes sont hautement spécifiques des réactions chimiques qu'elles catalysent et leur spécificité réside dans l'ajustement précis entre le substrat et le site actif. La figure 3-24 montre le site actif de la carboxypeptidase, une enzyme digestive qui clive les polypeptides des protéines des

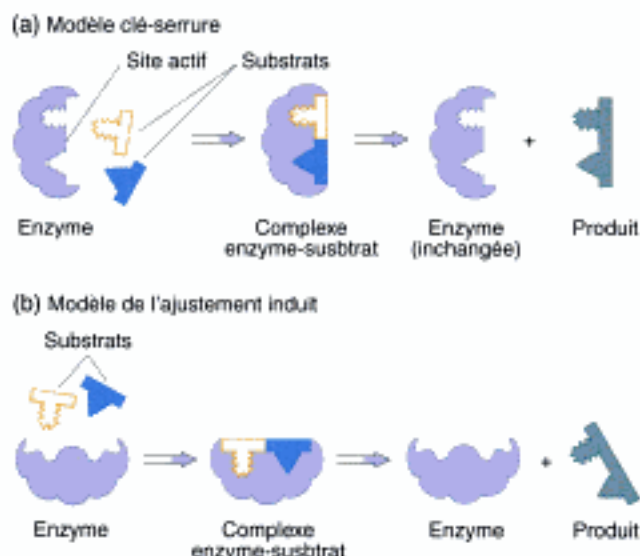


Figure 3-23 Représentation schématique de l'action d'une enzyme imaginaire qui assemble deux molécules substrats. (a) Dans le mécanisme « clé-serrure », les substrats ont une forme complémentaire de celle du site actif de l'enzyme. (b) Dans le modèle de l'ajustement induit, la fixation des substrats induit un changement conformationnel de l'enzyme.

aliments, entre les acides aminés glycine et tyrosine. Lorsqu'elle fixe son substrat (figure 3-24a), elle subit un changement de conformation qui favorise la catalyse.

Une part importante de la structure globulaire d'une enzyme est constituée de matériel non réactif qui maintient simplement la conformation du site actif. Cependant, la forme précise et les propriétés de liaison du site actif sont essentielles au fonctionnement correct d'une enzyme et un groupe d'acides aminés hautement spécifiques compose la structure du site actif. Les codons correspondant à ces acides aminés ont tendance à être regroupés dans le gène, mais en raison du repliement du polypeptide, certains codons peuvent être en-dehors du regroupement.

Dans le schéma suivant, les codons du site actif d'une protéine imaginaire sont indiqués en blanc et les autres codons de la protéine en noir.

MESSAGE

L'architecture d'une protéine est la clé de la fonction des gènes. La séquence spécifique d'acides aminés détermine la forme générale, les propriétés de liaison et la réactivité de la protéine.

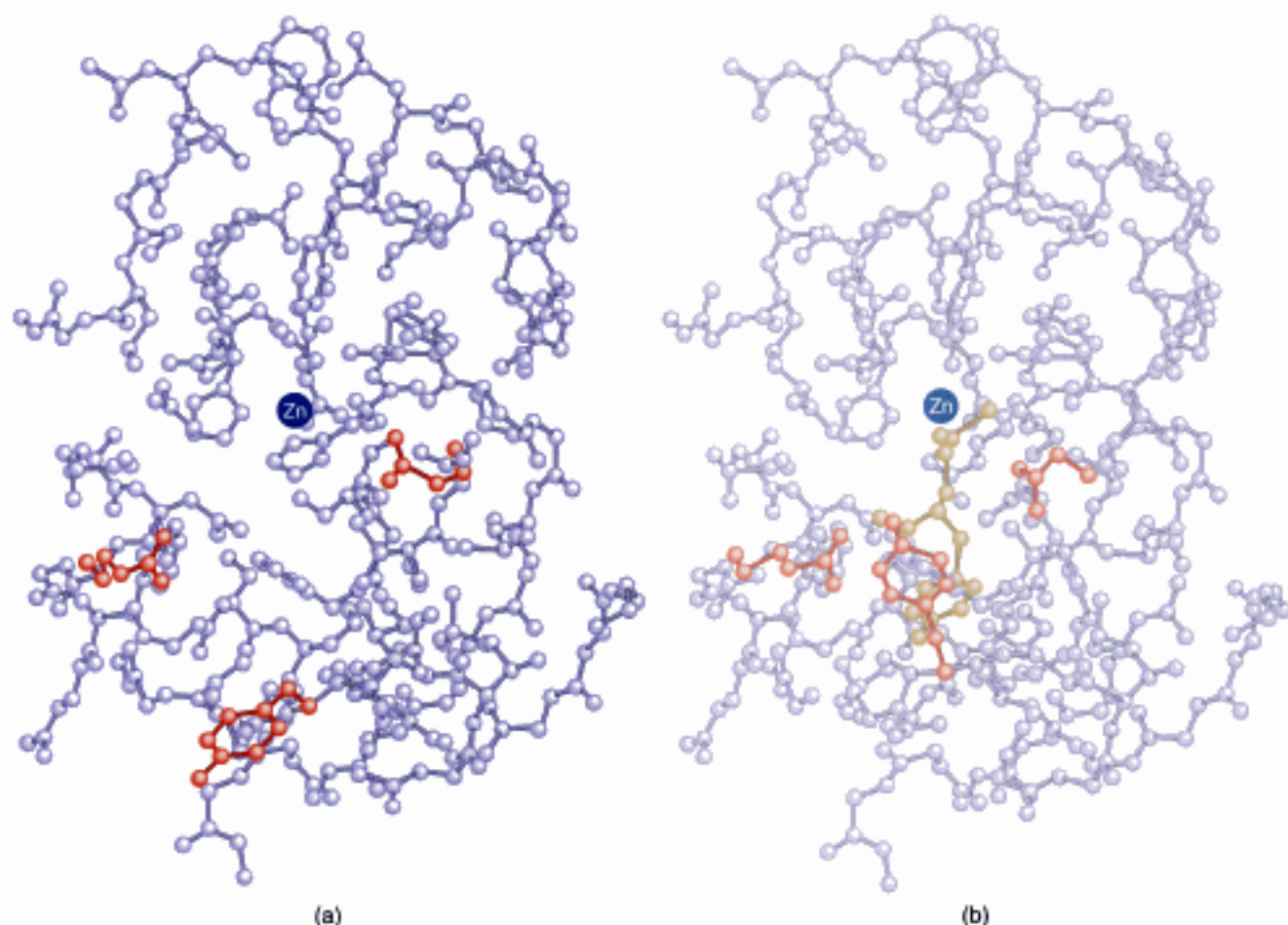
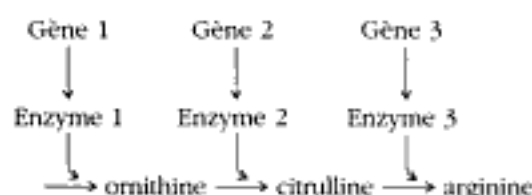


Figure 3-24 Le site actif de la carboxypeptidase, une enzyme du système digestif. (a) L'enzyme sans substrat. (b) L'enzyme avec son substrat (en jaune) au niveau du site actif. Trois acides aminés essentiels (en rouge) ont changé de position pour se placer plus près du substrat. La carboxypeptidase dégrade des protéines de l'alimentation. (D'après W. N. Lipscomb, *Proceedings of the Robert A. Welch Foundation Conferences on Chemical Research* 15, 1971, 140-141.)

Les voies enzymatiques. Les cellules synthétisent des substances cellulaires essentielles étape par étape, à partir de composés issus de l'environnement. Cette synthèse s'effectue par une succession de réactions faisant intervenir des enzymes, que l'on appelle des *voies biochimiques*. Par exemple, le champignon haploïde *Neurospora* synthétise sa propre arginine (un acide aminé) grâce à une série d'étapes biochimiques séquentielles. Les précurseurs immédiats dans cette voie métabolique sont l'ornithine et la citrulline comme le montre le schéma ci-dessous. L'ornithine est fabriquée à partir de précurseurs qui ne sont pas présentés ici. Elle est ensuite convertie en citrulline, qui est à son tour convertie en arginine. Ces conversions sont catalysées par des enzymes que l'on peut appeler enzymes 1 à 3, codées par trois gènes différents de synthèse de l'arginine que l'on peut nommer gènes 1 à 3 :



Des centaines de voies biochimiques mettant en jeu des milliers de réactions ont lieu à l'intérieur des cellules vivantes. Puisque la plupart de ces réactions sont catalysées par des enzymes, nous pouvons entrevoir le rôle de pivot que jouent les gènes dans le contrôle de la chimie de la vie.

Les allèles défectueux

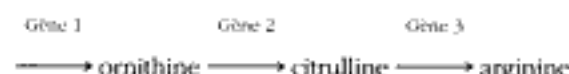
Les mutations et une petite leçon de nomenclature. En génétique, le phénotype standard d'un organisme est appelé **type sauvage**, car c'est le type observé dans la nature. Tous les gènes essentiels doivent être capables d'engendrer leurs produits fonctionnels pour donner lieu à ce phénotype « sauvage ». Ces allèles qui fonctionnent normalement sont appelés **allèles de type sauvage**.

Les **mutations**, qui sont des changements dans la séquence d'un gène, se produisent spontanément dans la nature. Puisque les allèles sont des formes des gènes, les mutations créent par définition de nouveaux allèles. En génétique, un gène est généralement symbolisé par une lettre ou par plusieurs lettres basées sur le mot qui décrit le phénotype produit par un allèle mutant. L'allèle de type sauvage correspondant est désigné par l'addition du signe (+) en exposant. Par exemple, le type sauvage de la couleur de l'œil de la mouche du vinaigre *Drosophila* est rouge, mais un allèle mutant présent sur le chromosome X produit des yeux blancs (white en anglais). L'allèle mutant est donc désigné par *w* et l'allèle sauvage par *w⁺*.

Comme nous l'avons vu précédemment, le champignon haploïde *Neurospora*, comme la plupart des

micro-organismes, peut synthétiser tous les composants cellulaires essentiels (tels que les acides aminés) à partir de substances inorganiques simples présentes dans le milieu de croissance. On dit donc que ces organismes sont **prototrophes** (ils se nourrissent de façon autonome). Certaines mutations rendent le champignon incapable de fabriquer des substances spécifiques. L'organisme mutant qui est alors formé est nommé **auxotrophe** (qui a besoin d'un apport extérieur de nourriture) car il doit se procurer certains composés essentiels dans l'environnement. Bien que ces mutations soient létales (mortelles) dans la nature, les souches mutantes peuvent être maintenues en vie en laboratoire par addition des substances manquantes dans leur milieu de croissance. Par exemple, certaines mutations rendent une souche incapable de produire sa propre arginine, acide aminé dont la synthèse a été décrite dans le paragraphe précédent. Un allèle mutant est désigné par *arg* (qui signifie « a besoin d'arginine ») et l'allèle sauvage, qui permet la synthèse de l'arginine est *arg⁺*. Parfois le symbole + est utilisé seul lorsque sa signification est évidente. Ainsi dans l'exemple de *Neurospora* les deux allèles deviendraient *arg* et + sous-entendant que + correspond à l'allèle de type sauvage de ce gène.

Un auxotrophe *arg* se développera si l'on ajoute de l'arginine au milieu de croissance. Rappelez-vous que la dernière portion de la voie de biosynthèse de l'arginine est



Une mutation auxotrophe *arg* pourrait affecter n'importe lequel des trois gènes contrôlant ces conversions. On peut distinguer les trois possibilités car un mutant bloqué dans la dernière étape (contrôlée par le gène 3) ne poussera que s'il y a supplémentation en arginine. Un mutant bloqué dans l'étape intermédiaire (gène 2), ne poussera que si on lui ajoute de l'arginine ou de la citrulline et un mutant bloqué dans la première étape (gène 1), seulement si on lui ajoute de l'arginine, de la citrulline ou de l'ornithine. Cette logique a été utilisée dans la découverte de cette voie métabolique, comme le décrit l'encadré Les grandes découvertes de la génétique 3-1. Nous pourrions appeler ces trois gènes (en utilisant leurs allèles sauvages) *arg-1⁺*, *arg-2⁺* et *arg-3⁺*.

Dans certaines situations, les généticiens utilisent des lettres majuscules et minuscules pour désigner les allèles, telles que *A* et *a*. Dans ce symbolisme, le type sauvage n'est pas mentionné. Cela convient aux cas pour lesquels on ne connaît pas l'allèle sauvage, par exemple celui des lignées végétales ou animales ayant subi de nombreux croisements. Cette écriture est également utile pour dénommer les allèles qui déterminent les morphes d'un polymorphisme, puisque tous les morphes sont de type sauvage.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 3-1

Beadle et Tatum et la nature de l'action du gène

En 1908, le médecin anglais Archibald Garrod suggéra que certaines maladies humaines sont provoquées par des « erreurs héréditaires de métabolisme ». Par exemple l'alcaptonurie (« maladie de l'urine noire »), une maladie récessive, est caractérisée par l'élimination d'acide homogentisique dans l'urine, qui devient noire lorsqu'elle est exposée à l'air. Garrod supposa que cette élimination inhabituelle était due à l'absence d'une enzyme chez les malades. Cette enzyme décompose normalement l'acide homogentisique. On observe chez les malades une augmentation de la concentration de cet acide dans les tissus et son élimination. Ce travail résultait de l'idée selon laquelle l'action des gènes consiste à réguler enzymatiquement les réactions chimiques de la cellule. La même idée apparut plusieurs fois dans la littérature au cours des trois décennies suivantes, mais c'est la recherche de George Beadle et de Edward Tatum sur le champignon *Neurospora* qui permit d'exposer clairement ce mode d'action des gènes.



Spores sexuelles de *Neurospora*

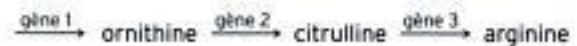
Beadle relata plus tard le souvenir suivant :

En 1940, nous décidâmes de passer de *Drosophila* à *Neurospora*. Cela se fit de la manière suivante. Tatum donnait des cours de génétique biochimique et j'y assistai. En écoutant l'un d'eux — ou peut-être en n'écoutant pas aussi bien que je l'aurais dû — il me vint soudain à l'esprit qu'il devait être possible d'inverser la procédure que nous avions suivie et au lieu d'essayer d'élucider l'origine chimique de différences génétiques connues, nous devrions être capables de sélectionner des mutants chez lesquels des réactions chimiques connues étaient bloquées. *Neurospora* constituait un organisme de choix sur lequel tenter cette approche, car son cycle vital avait été établi par Dodge et Lindgren et ce champignon pouvait probablement être cultivé sur un milieu de culture de composition connue. L'idée consistait à sélectionner des mutants incapables de synthétiser des métabolites connus tels que des vitamines et des acides aminés, que l'on pourrait ajouter au milieu. De cette façon, un mutant incapable de fabriquer une vitamine donnée pourrait croître en présence de cette vitamine et être classifié sur la base de cette réponse différente en milieu enrichi, ou dépourvu, de cette vitamine.

Effectivement, des mutants présentant une déficience nutritionnelle purent facilement être obtenus en irradiant des cellules. Ces mutants ne se développaient que si le milieu était enrichi en l'une des nombreuses substances cellulaires essentielles (par exemple l'acide aminé arginine), indiquant que la synthèse de cette substance était déficiente chez le mutant. En outre, des précurseurs connus ou présumés de l'arginine purent être testés, comme l'ornithine ou la citrulline. Les mutants nécessitant de l'arginine ont alors été classifiés en trois groupes selon leur réponse à l'ornithine et à la citrulline :

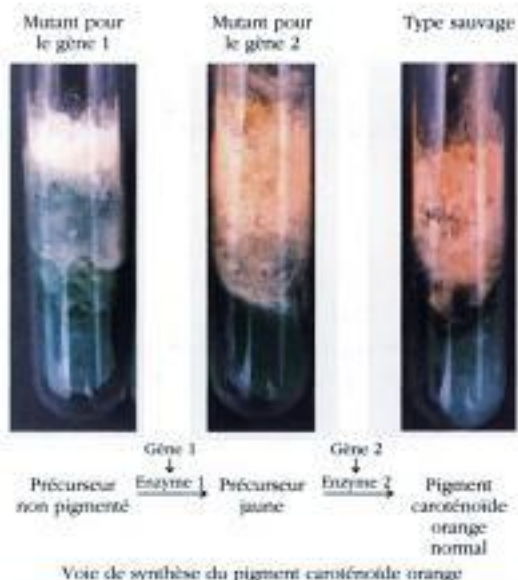
	Groupe 1	Groupe 2	Groupe 3
Ornithine	Croissance	Aucune	Aucune
Citrulline	Croissance	Croissance	Aucune
Arginine	Croissance	Croissance	Croissance

Beadle, Tatum et leurs étudiants déduisirent que les groupes 1 à 3 devaient avoir chacun une mutation différente dans l'un des trois gènes agissant successivement pour fabriquer le produit final :



Seules les substances qui étaient produites normalement après l'étape inactivée par la mutation pouvaient permettre la croissance des mutants.

À partir de ces expériences Beadle et Tatum émisent l'hypothèse selon laquelle un gène agit généralement en déterminant une activité enzymatique spécifique — l'hypothèse « un gène — une enzyme ». Des travaux ultérieurs montrèrent que cette formule décisive, qui soulignait le lien naissant entre la biochimie et la génétique, devait être énoncée plus précisément « un gène — un polypeptide ».



Les types de mutations. Les allèles mutants les plus simples présentent une séquence d'ADN modifiée en une seule position du gène. De tels allèles mutants sont composés d'une séquence nucléotidique normale exceptée dans la zone de changement identifiée sous le nom de **site mutant**. En raison de la façon dont le code génétique est lu au cours de la traduction, une mutation qui change une seule paire de bases en une autre entraîne souvent le remplacement d'exactly un acide aminé par un autre. En raison de la colinéarité des gènes et de la structure primaire des protéines, les mutations de la séquence nucléotidique de l'ADN modifient la chaîne d'acides aminés à la même position relative (voir Les grandes découvertes de la génétique 3-2). Un changement touchant un seul nucléotide dans un codon au niveau d'un site fonctionnel essentiel de la protéine (tel que le site actif d'une enzyme) peut conduire au remplacement d'un acide aminé aboutissant à une fonction aberrante. Par exemple, un changement dans la paire de nucléotides C•G → T•A peut conduire au remplacement de la proline par la leucine, un acide aminé aux propriétés très différentes, ce qui risque d'entraîner des effets néfastes sur la fonction de la protéine.

	Type sauvage	Mutant
ADN	5'CCG 3'	5'CTG 3'
	3' GGC 5'	3'GAC 5'
ARNm	5'CCG 3'	5'CTG 3'
Acide aminé	Proline	Leucine

L'addition ou la délétion d'une seule paire de nucléotides est un autre type courant de mutation simple. On les appelle des *mutations par décalage du cadre de lecture* car elles ne permettent plus une lecture correcte du message. En effet, il y a un décalage d'un nucléotide dans chaque acide aminé situé « en aval » du site mutant dans la traduction. Considérons par exemple les conséquences d'une délétion d'un nucléotide dans le segment d'ARNm suivant.

	Délétion
Type sauvage	5'...UUG CUG AGG CCC GAG U.....3'
Délétion	5'...UUG CGA GGC CCG AGU3'

Tous les codons en aval (c'est-à-dire du côté 3') de la délétion seront traduits de façon incorrecte et la plupart des acides aminés correspondants seront incompatibles avec la fonction de la protéine.

Les conséquences des mutations sur la fonction cellulaire. Généralement, les mutations entraînent une fonction protéique réduite ou nulle. Une mutation avec une fonction réduite est appelée **mutation partielle** (*leaky*) car une partie de la fonction de type sauvage transparait dans le phénotype. Une mutation qui supprime la fonction protéique est appelée **mutation complète**

(*null*). Les changements qui n'affectent pas la fonction d'une protéine sont des **mutations silencieuses**.

Certaines mutations affectent le processus de transfert de l'information sans modifier directement le code génétique. Par exemple, certaines mutations entraînent un dysfonctionnement non par un effet sur la séquence d'acides aminés, mais par altération de l'épissage des introns. Puisque l'épissage des introns dépend de séquences nucléotidiques spécifiques situées aux limites exon-intron et à l'intérieur de l'intron, si ces sites sont mutés, l'intron ne peut être excisé et aucun ARNm fonctionnel ne sera produit. Une mutation similaire altère la régulation du gène. Par exemple, la séquence promotrice sur laquelle se fixe l'ARN polymérase est essentielle, et si un changement se produit dans cette séquence, le gène peut ne pas être transcrit du tout ou être transcrit à des taux anormalement faibles (ou élevés). Enfin, des mutations dans la queue 3' non traduite du transcrit peuvent avoir des effets marqués sur la stabilité d'un ARNm.

Des exemples de sites mutants ainsi que leurs conséquences phénotypiques sont présentés dans la figure 3-25 à la page 74.

MESSAGE

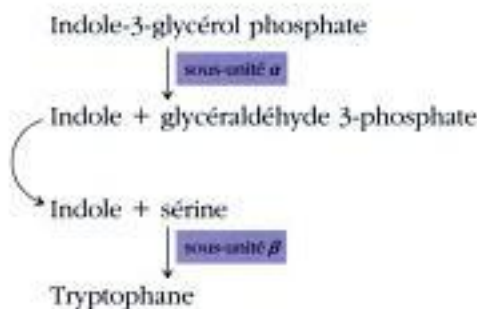
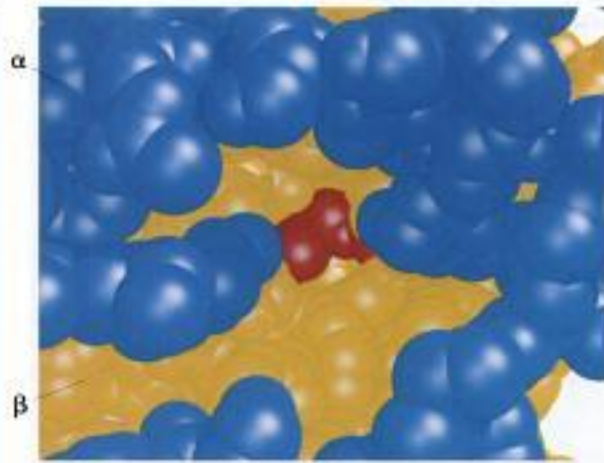
Des mutations peuvent conduire à un dysfonctionnement des gènes par des changements dans les séquences qui codent les protéines ou qui sont importantes pour le traitement de l'information.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 3-2

Charles Yanofsky et la colinéarité du gène et de la structure protéique

En 1941, George Beadle et Edward Tatum énoncèrent l'hypothèse un gène-une enzyme. Lorsque James Watson et Francis Crick élucidèrent la structure de l'ADN en 1953, il semblait probable qu'il existât une correspondance linéaire entre la séquence nucléotidique de l'ADN et la séquence d'acides aminés de la protéine (par exemple une enzyme). Cependant, il fallut attendre 1963 pour obtenir une démonstration expérimentale de cette colinéarité, par deux groupes de recherche, dont l'un était dirigé par Charles Yanofsky à l'Université de Stanford.

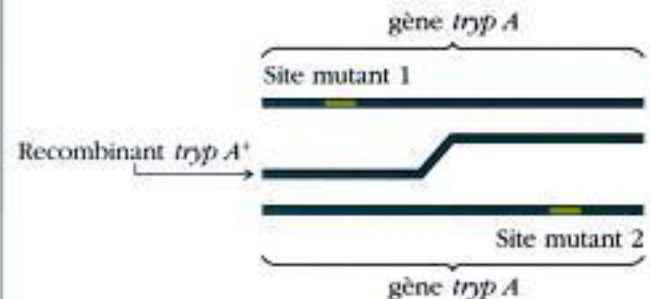
Yanofsky avait induit 16 allèles mutants du gène *tryp A* d'*Escherichia coli* dont l'allèle de type sauvage codait la sous-unité α de l'enzyme tryptophane synthétase. La tryptophane synthétase est un tétramère composé de deux sous-unités α et de deux sous-unités β . Elle catalyse les réactions suivantes :



On sait maintenant que les sous-unités agissent de façon séquentielle, tout en maintenant l'indole.

Yanofsky montra que les allèles mutants de *trpA* qui produisaient une enzyme inactive étaient globalement semblables sur le plan structural à l'allèle sauvage, mais qu'ils en différaient au niveau de 16 sites mutants distincts. Il parvint à cartographier la position des sites mutants. En 1963, le séquençage de l'ADN n'avait pas été inventé, mais les sites mutants purent être cartographiés par une analyse par crossing-over (Chapitre 5). Plus deux sites étaient proches, plus leur fréquence de recombinaison aboutissant à des recombinants de type sauvage était

faible. Dans l'exemple suivant, le crossing-over de deux chromosomes homologues se produit entre l'allèle mutant *trpA-1* sur un chromosome et *trpA-2* sur l'autre chromosome. Le chromosome résultant (présenté au centre) combine les « bonnes » parties de chaque gène mutant pour produire un gène sans mutation (un type sauvage).

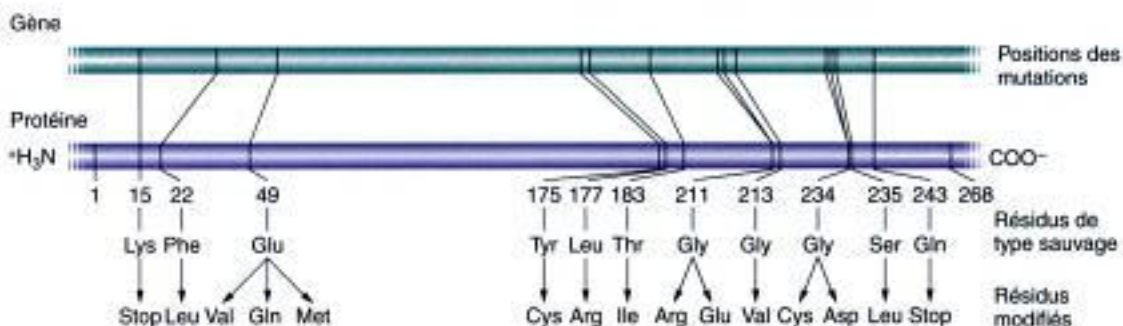


Grâce à l'analyse biochimique de la protéine *trpA*, il fut montré que chacun des mutants résultait d'une substitution d'acide aminés en différentes positions de la protéine. En comparant les positions relatives des sites mutants dans le gène avec les positions relatives des changements d'acides aminés dans la protéine, leur collinéarité fut établie.

Dans le résumé de son article de 1966, Yanofsky écrit :

Le concept de collinéarité entre la structure des gènes et la structure des protéines a été examiné chez 16 mutants présentant des modifications dans un segment du gène A et de la protéine A [que l'on appelle à présent α] de la tryptophane synthase. Les résultats obtenus démontrent une correspondance linéaire entre les deux structures et montrent aussi que les valeurs de recombinaison génétique sont représentatives des distances entre les acides aminés dans la protéine correspondante.

Il a ensuite été montré que ces résultats pouvaient s'appliquer de manière générale aux mutations dans d'autres protéines.



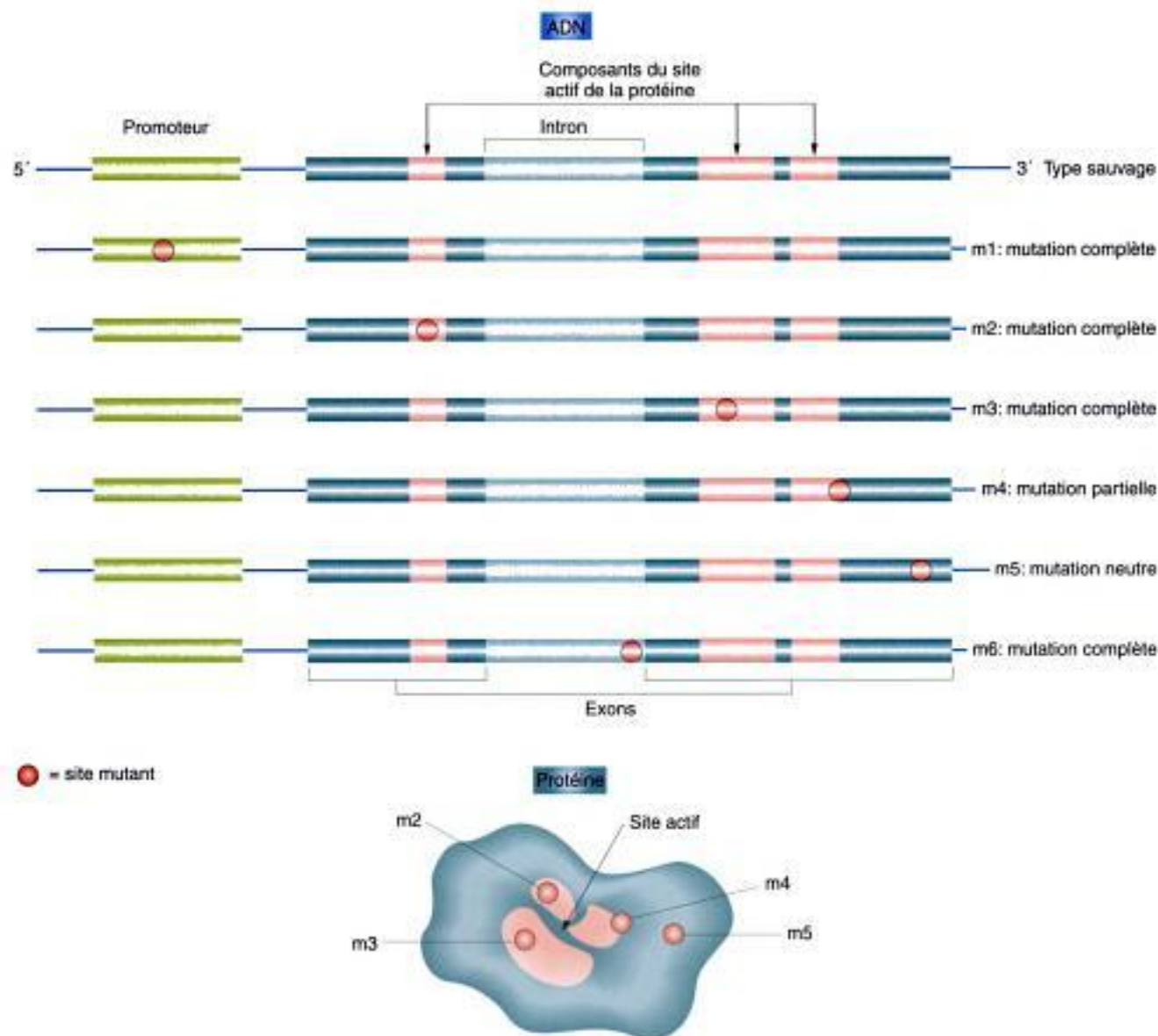


Figure 3-25 Positions des sites mutants avec leurs conséquences fonctionnelles.

Les maladies génétiques. L'une des origines courantes des maladies génétiques chez l'homme est une déficience enzymatique résultant d'une mutation. Les cellules humaines comportent deux jeux de chromosomes ; ainsi, tous les gènes sont représentés deux fois (paires de gènes). Normalement, les deux exemplaires de chaque paire sont des allèles de type sauvage. Cependant, les individus qui portent une paire d'allèles défectueux d'un gène codant une enzyme, présenteront une activité enzymatique réduite ou nulle et exprimeront les symptômes

d'une maladie. Le tableau 3-4 présente quelques exemples et donne une idée de l'ampleur de l'implication de la génétique dans les maladies humaines. La figure 3-26 à la page 76 présente un fragment de la carte du métabolisme humain. Il illustre la façon dont un groupe de maladies, dont certaines nous sont familières, peuvent résulter du blocage d'étapes adjacentes dans les voies métaboliques. Considérons la phénylcétonurie (PCU), une maladie présentée dans la figure 3-26. Un défaut dans l'enzyme phénylalanine hydroxylase entraîne une

TABLEAU 3-4

Exemples représentatifs d'enzymopathies : maladies héréditaires chez l'homme pour lesquelles l'activité modifiée (en général une déficience) d'une enzyme spécifique a été démontrée

Maladie	Enzyme dont l'activité est déficiente*	Maladie	Enzyme dont l'activité est déficiente*
Acatalasie	Catalase	Hydroxyprolinémie	Hydroxyproline oxydase
Acidurie argininosuccinique	Argininosuccinase	Hyperlysinémie	Lysine-cétoglutarate réductase
Albinisme	Tyrosinase	Hypophosphatasie	Phosphatase alcaline
Alcaptonurie	Oxydase acide homogentisique	Immunodéficiences	Adénosine désaminase
Angiokératome, diffus (maladie de Fabry)	Céramide trihexosidase		Uridine monophosphate kinase
Apnée médicamenteuse	Pseudocholinestérase	Lipogranulomatoses de Farber	Céramidase
Argininémie	Arginase	Maladie de Gaucher	Glucocérébrosidase
Ataxie intermittente	Pyruvate décarboxylase	Maladie de Krabbe	Galactosylcéramide β -Galactosidase
Citrullinémie	Synthétase acide argininosuccinique	Maladie de l'urine du sirop d'érable	Décarboxylase cétoacide
Cystathioninurie	Cystathionase	Maladie de Niemann-Pick	Sphingomyélinase
Déficit en aldostérone	18-Hydroxydésydrogénase	Maladie de Refsum	Oxydase acide phytanique
Déficit en phosphatase acide	Phosphatase acide	Maladie de Sandhoff (gangliosidose GM2, type II)	Hexosaminidase A,B
Encéphalomyélopathie nécrosante de Leigh	Pyruvate carboxylase	Maladie de Tay-Sachs	Hexosaminidase A
Galactosémie	Galactose 1-phosphate uridylyltransférase	Maladie de Wolman	Lipase acide
Gangliosidose, GM ₁ : généralisée, de type I ou forme infantile	β -Galactosidase A, B, C	Ornithinémie	Ornithine-cétoacide amino transférase
Gangliosidose, GM ₁ : de type II ou forme juvénile	β -Galactosidase B, C	Pentosurie	Xylitol déshydrogénase (L-xylulose réductase)
Goutte	Hypoxanthine-guanine phosphoribosyl-transférase Phosphoribosyl pyrophosphate (PRPP) synthétase (activité accrue)	Phénylcétonurie	Phénylalanine hydroxylase
		Syndrome d'Ehlers-Danlos, type V	Lysyl oxydase
		Syndrome de Crigler-Najjar	Glucuronyltransférase
		Syndrome de Richner-Hanhart	Tyrosine aminotransférase
		Xeroderma pigmentosum	Endonucléase UV spécifique
Granulomatose chronique	Oxydase du nicotinamide adénine dinucléotide phosphate réduit (NADPH)		

Source: Victor A. McKusick, *Mendelian Inheritance in Man*, 4^e éd. Copyright © 1975 by Johns Hopkins University Press.

* La forme de goutte liée à une activité accrue de PRPP est la seule maladie de cette liste qui soit causée par une augmentation d'activité enzymatique.

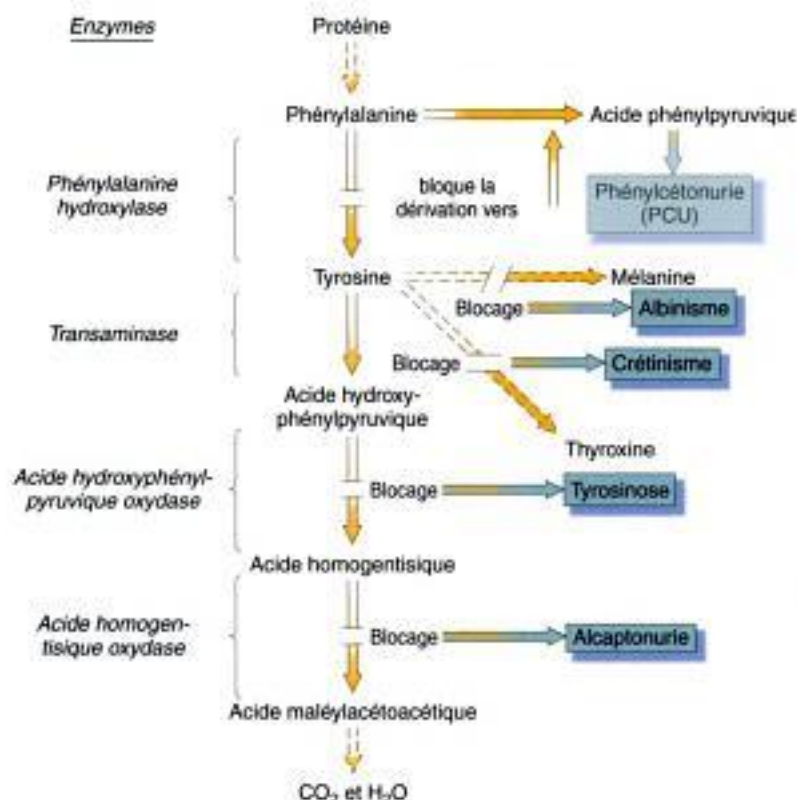


Figure 3-26 Petite portion de la carte du métabolisme humain qui montre les conséquences de différentes déficiences enzymatiques spécifiques. (Les phénotypes des maladies sont indiqués dans les cases colorées). (D'après I. M. Lerner et W. J. Libby, *Heredity, Evolution and Society*, 2^e éd. Copyright © 1976 par W. H. Freeman et Company.)

accumulation de phénylalanine (provenant des protéines de l'alimentation). À des concentrations élevées, la phénylalanine est transformée en acide phénylpyruvique, une substance qui interfère avec le développement du système nerveux, provoquant un retard mental chez un enfant possédant deux copies de l'allèle défectueux. Si la concentration élevée d'acide phénylpyruvique est détectée peu de temps après la naissance, le bébé peut être placé sous un régime spécial, à faible teneur en phénylalanine. Il se développera alors sans présenter de retard.

Revenons à la mucoviscidose que nous avons évoquée au début de ce chapitre. Elle n'est pas provoquée par une déficience enzymatique, mais par l'inactivité d'une protéine qui contrôle le passage des ions chlorure à travers des membranes dans les tissus sécrétoires (figure 3-27). La mutation la plus courante est une délétion de trois nucléotides, qui conduit au retrait de l'acide aminé phénylalanine de la séquence primaire de la protéine, comme le montre la figure 3-28. La position de la délétion attendue dans la protéine est indiquée dans la figure 3-27. Cependant, chez les malades atteints de mucoviscidose, on ne trouve jamais dans la membrane de protéines déficientes dépourvues de phénylalanine à la position de la délétion. La cellule semble capable de détecter et de détruire la protéine défectueuse. Quoi qu'il en soit, l'absence de cette protéine et la perturbation de l'équilibre des ions chlorure et des ions sodium sont responsables de la formation du mucus épais qui provoque les symptômes respiratoires ainsi que d'autres problèmes.

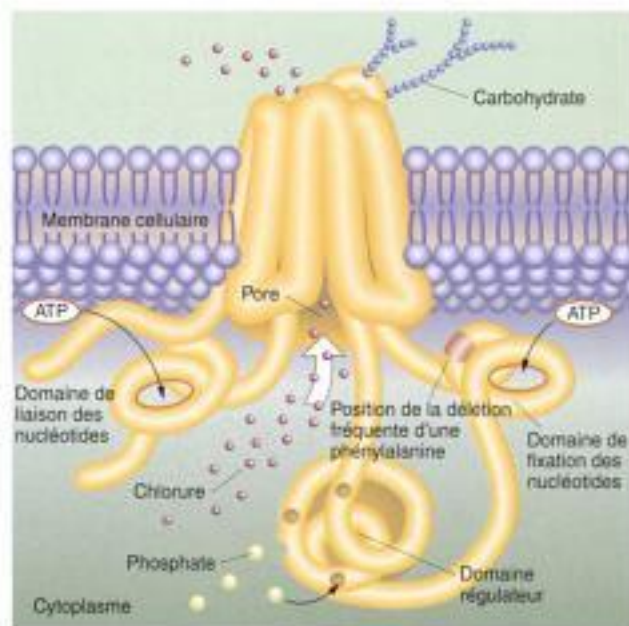


Figure 3-27 La protéine de la mucoviscidose (CF : cystic fibrosis en anglais) qui agit comme un canal perméable au chlore dans la membrane cellulaire externe. Le site de la délétion de trois paires de nucléotides, fréquente parmi les allèles mutants de CF est indiqué.

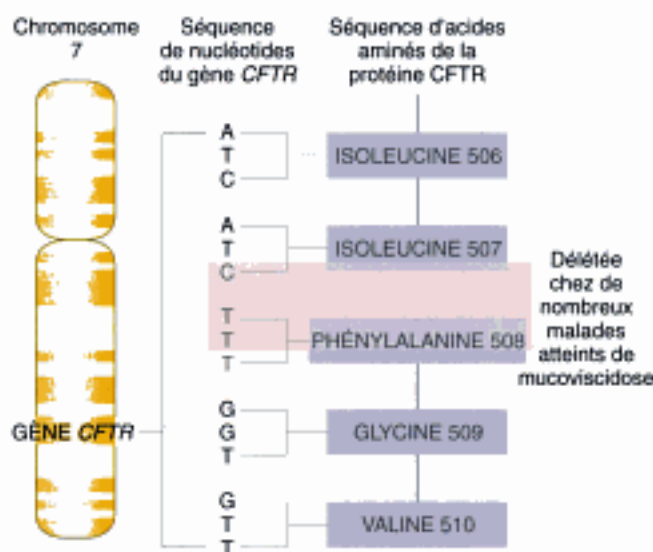


Figure 3-28 Position du gène de la mucoviscidose sur le chromosome 7, avec un agrandissement montrant la nature de la déletion fréquente de 3 pb qui enlève une phénylalanine de la séquence polypeptidique.

Ndr : Pour une description précise et détaillée du défaut moléculaire, voir les trois articles du travail original publié dans *Science*, 1989, 245, pp. 1059-1080.

PROTÉINES DÉFICIENTES, DOMINANCE ET RÉCESSIVITÉ

La mucoviscidose et la phénylcétonurie sont toutes deux des maladies **récessives**. Ceci signifie que les deux copies du gène doivent être déficientes pour que la maladie soit exprimée. Chez une personne portant un allèle normal sur un chromosome et un allèle défectueux sur le chromosome homologue, la présence de cet allèle normal produit suffisamment de protéine pour assurer la fonction normale dans la cellule. Dans ce cas, le phénotype normal est dit **dominant** car l'allèle normal est exprimé en présence de l'allèle anormal et semble le dominer. Même si les termes *dominant* et *récessif* sont réservés exclusivement à la description des phénotypes, ils sont souvent appliqués aux différents allèles responsables. Puisqu'une « dose » haploïde de l'allèle normal fournit suffisamment de protéine pour permettre un fonctionnement normal, cette situation générale est qualifiée d'**haplo-suffisance**. On attribue aux allèles récessifs des symboles en lettres minuscules. Les trois génotypes possibles pour un cas général d'haplo-suffisance sont représentés dans le tableau suivant, qui montre les relations de dominance. Dans le tableau, *a* désigne l'allèle mutant (déficient). [Dans les génotypes, remarquez la barre oblique (/) entre les paires d'allèles. Cette barre oblique est utilisée pour indiquer que l'on considère les allèles situés à la même position (locus) sur les deux chromosomes homologues dans les organismes diploïdes, et nous verrons des extensions de cet usage dans les chapitres suivants.]

Relations de dominance dans le cas de gènes présentant une haplo-suffisance

Génotype	Nom du génotype	Phénotype
a'/a'	Homozygote dominant	Normal
a'/a	Hétérozygote	Normal
a/a	Homozygote récessif	Déficient

Le tableau montre que le phénotype d'un hétérozygote (un individu possédant deux allèles différents pour un gène) est un bon test de dominance ; si l'hétérozygote présente un phénotype normal, alors le phénotype représenté par l'allèle mutant doit être récessif. La figure 3-29 à la page suivante montre le fonctionnement de ce système au niveau moléculaire.

Cependant, chez les hétérozygotes de certains gènes, l'allèle normal *ne peut* fournir suffisamment de produit protéique pour remplir la fonction cellulaire normale ; on qualifie cette situation d'**haplo-insuffisance**. Dans ce cas, c'est l'allèle déficient (mutant) qui est dominant. On utilise des lettres majuscules pour désigner les mutations dominantes. Comme pour les allèles récessifs, l'allèle de type sauvage est indiqué par un signe plus placé en exposant après la lettre. Le cas général de l'haplo-insuffisance peut être écrit de la façon suivante, où *B* est un allèle mutant :

Relations de dominance dans le cas de gènes présentant une haplo-insuffisance

Génotype	Nom du génotype	Phénotype
B'/B'	Homozygote récessif	Normal
B'/B	Hétérozygote	Normal
B/B	Homozygote dominant	Déficient

Une fois encore, nous voyons que le test de dominance est le phénotype de l'hétérozygote ; ici, le phénotype mutant (déficient) est clairement dominant. Il existe d'autres mécanismes possibles où les allèles mutants peuvent exprimer une dominance — par exemple en adoptant une fonction nouvelle — comme nous le verrons dans les chapitres 7 et 14. On les distingue des dominants haplo-insuffisants et on les appelle des *dominants par gain-de-fonction*.

D'une manière générale, il existe bien plus de gènes haplo-suffisants que de gènes haplo-insuffisants, mais cela dépend cependant beaucoup de l'organisme et du contexte génétique.

MESSAGE

Chez un hétérozygote normal/mutant, si l'unique allèle normal peut fournir suffisamment de protéine fonctionnelle, l'allèle mutant est récessif. Dans le cas contraire, l'allèle mutant est dominant.

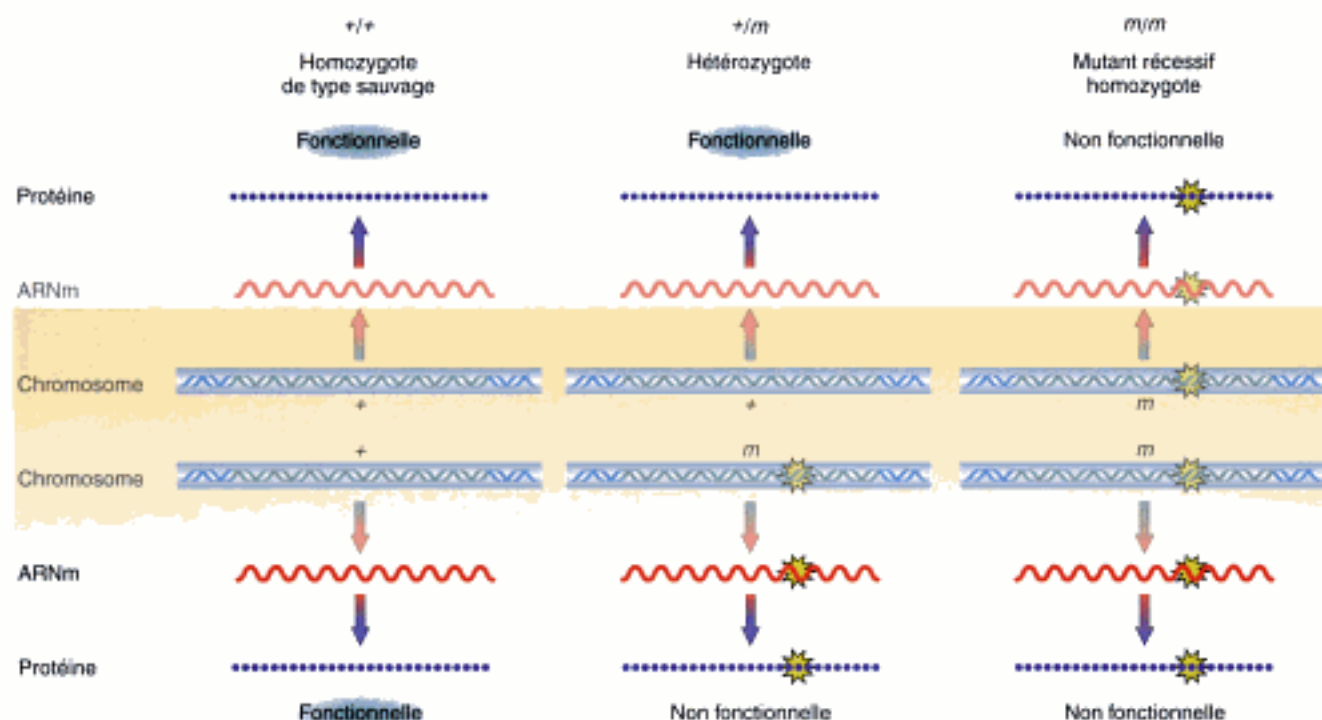


Figure 3-29 Relations de dominance dans le cas de gènes haplo-suffisants. Chez un hétérozygote, même si la copie mutée du gène produit une protéine non fonctionnelle, la copie de type sauvage produit suffisamment de protéine fonctionnelle pour conduire au phénotype sauvage.

LA DIVISION FONCTIONNELLE DES TÂCHES DANS L'ENSEMBLE DES GÈNES

Quels sont les différents types de gènes nécessaires pour construire un organisme ? On ne peut donner de réponse complète à cette question au stade actuel des recherches, mais on a une idée des catégories générales de gènes et des tailles relatives de ces catégories. La séquence nucléotidique complète du génome de la levure de bière *Saccharomyces cerevisiae* est connue (c'est l'un des rares génomes eucaryotes intégralement séquencés jusqu'à présent). Ce séquençage a permis de déterminer que ce génome contient environ 6340 gènes. Parmi ceux-ci, 140 codent des ARNr, que l'on trouve sous la forme de répétitions en tandem sur le chromosome 12, 40 codent de petits ARN nucléaires, disséminés sur les chromosomes et 275 codent des ARNt (répartis en 45 + familles + — voir chapitre 12). Les 5885 gènes restants codent des protéines. On ne sait toujours pas si ce pourcentage d'ARN non traduit et de gènes codant des protéines est typique de tous les Eucaryotes.

L'ensemble des gènes codant les protéines d'un organisme est appelé **protéome**. La répartition des tâches à l'intérieur du protéome est un sujet d'intérêt considérable pour la compréhension des types de gènes nécessaires au fonctionnement d'un organisme. D'après les gènes humains étudiés jusqu'ici, une estimation des

pourcentages des différentes fonctions protéiques a pu être établie :

Fonction	Pourcentage
Processus de réplication de l'ADN, de transcription et de traduction	22
Métabolisme	17
Division cellulaire	12
Défense	12
Communications intra- et intercellulaires et régulation des gènes	12
Fonction inconnue	8
Structure	17

En ce qui concerne les enzymes mentionnées plus haut dans ce texte, l'ARN polymérase et l'aminoacyl-ARNt synthétase se trouvent dans la première catégorie. Les enzymes contrôlant la biosynthèse de l'arginine chez *Neurospora* et celles impliquées dans la PCU et la mucoviscidose chez l'homme se trouvent dans la deuxième catégorie.

Le séquençage d'une grande partie du génome de la plante modèle *Arabidopsis thaliana*, a permis de se faire une idée de la répartition du protéome. Celle-ci est schématisée dans la figure 3-30.

À partir de ces estimations, les contours de l'architecture génétique de la vie sur cette planète ont été dessinés, y compris ceux de l'espèce humaine.

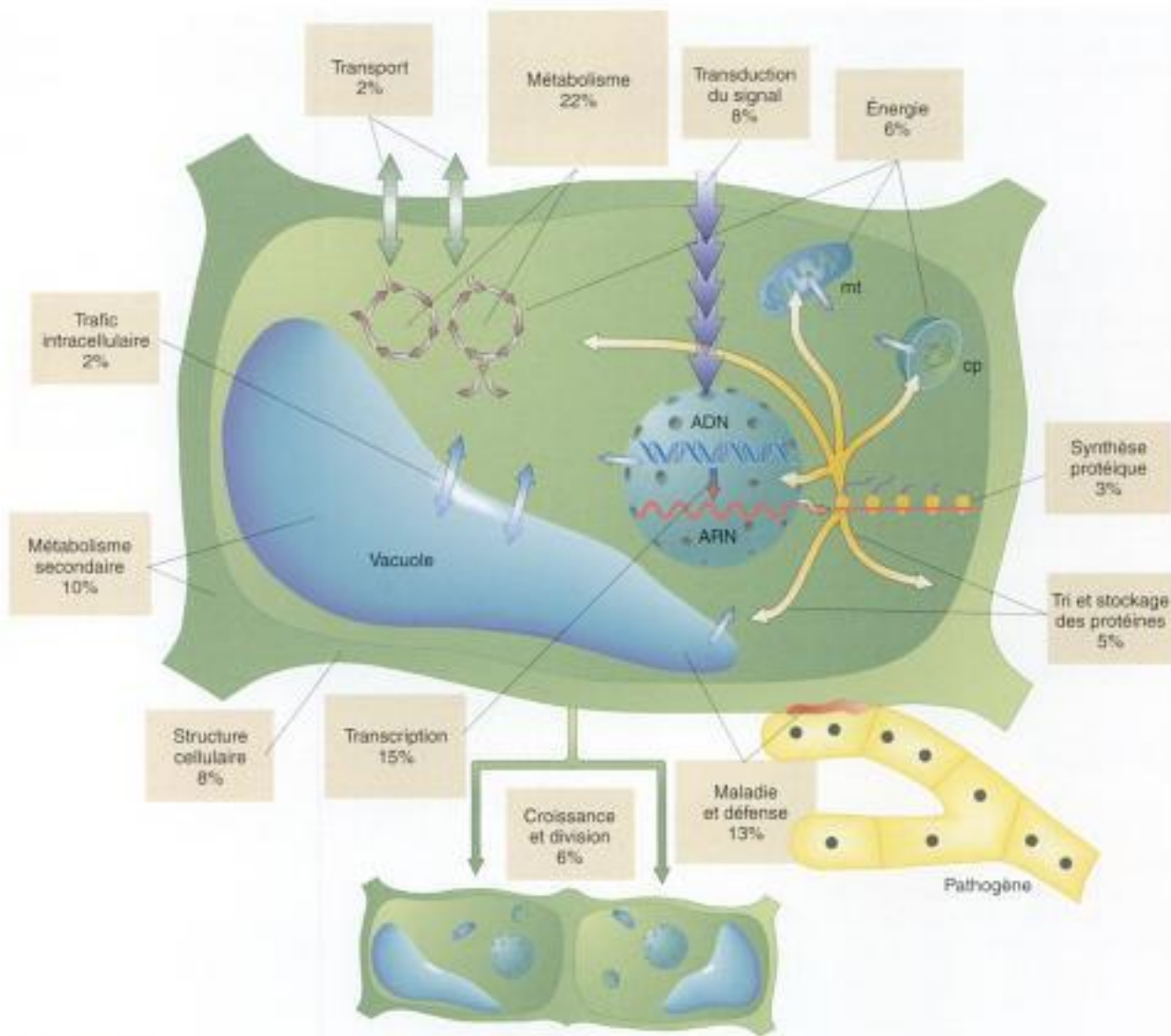


Figure 3-30 Les proportions des différentes catégories de gènes codant des protéines, estimées d'après celles des gènes connus actuellement chez la plante *Arabidopsis thaliana*.

RÉSUMÉ

Le transfert de l'information dans les cellules se fait par la transcription de l'ADN en ARN et la traduction de l'ARN en protéines. Les fonctions de l'ADN et de l'ARN sont basées sur deux grands principes : la complémentarité des bases entre les acides nucléiques simple-brin et la fixation de diverses protéines de liaison à l'ADN et à l'ARN. La plupart des gènes codent des protéines et leurs transcrits sont des ARNm ou des pré-ARNm. Toutefois, une minorité de gènes codent des ARN fonctionnels tels que les ARNt et les ARNr qui ne sont jamais traduits. L'ARN est une molécule simple-brin qui possède une séquence nucléotidique complémentaire de celle du brin d'ADN à partir de laquelle il a été transcrit. Pour chaque gène, un seul brin d'ADN sert de matrice pour la transcription. Ce brin est orienté dans le sens 3'→5'. La synthèse du transcrit

d'ARN commence par l'extrémité 5' du transcrit et se déroule dans le sens 5'→3'.

Les molécules d'ARNm eucaryotes possèdent une coiffe à leur extrémité 5' et une queue de poly(A) ajoutée à leur extrémité 3'. De nombreux gènes eucaryotes possèdent des introns interrompant la séquence codante. Ils sont excisés du transcrit avant la traduction.

La fonction des protéines est déterminée par leur structure primaire (leur séquence d'acides aminés) qui est déterminée par la séquence nucléotidique. La séquence nucléotidique de l'ARNm est traduite trois nucléotides (un codon) à la fois, dans le sens 5'→3'. Chaque codon correspond à un acide aminé spécifique. Chaque acide aminé est associé à un ARNt spécifique, qui se place sur le ribosome, face au codon approprié de l'ARNm. Des

changements dans la séquence nucléotidique de l'ADN produisent une séquence d'acides aminés modifiée, ce qui conduit généralement à la perte de la fonction protéique.

Un allèle déficient d'un gène codant une protéine est récessif si, en présence de l'allèle normal chez un hétéro-

zygote, l'allèle normal peut fournir assez de protéine pour subvenir aux besoins de la cellule (haplo-suffisance). Si ce n'est pas le cas (haplo-insuffisance), l'allèle déficient est dominant.

DIAGRAMME CONCEPTUEL

Dessinez un diagramme conceptuel mettant en relation le maximum de termes parmi ceux qui sont proposés. Remarquez que ces termes ne sont pas écrits dans un ordre particulier.

gène / transcription / traduction / ARN polymérase / dominant / ribosome / protéine / mutation / haplo-insuffisance

PROBLÈME RÉSOLU

1. En utilisant le dictionnaire des codons, indiquez les conséquences sur la traduction, de l'addition d'une base supplémentaire d'adénine au début de la séquence codante suivante.

–CGA–UCG–GAA–CCA–CGU–GAU–AAG–CAU–
– Arg – Ser – Glu – Pro – Arg – Asp – Lys – His –

SOLUTION

Avec l'addition d'un A au début de la séquence codante, le cadre de lecture est décalé et un groupe d'acides aminés différent est spécifié par la séquence, comme on le voit ici. (Remarquez que plusieurs codons stop sont apparus, le premier, UAA, suffit à lui seul à provoquer la terminaison de la chaîne.)

–ACG–AUC–GGA–ACC–ACG–UGA–UAA–GCA–
– Thr – Ile – Gly – Thr – Thr – stop – stop –

PROBLÈME RÉSOLU

2. La levure (un haploïde) fabrique normalement sa propre proline. Six mutants auxotrophes ont été obtenus. Ils ne fabriquent pas de proline et ont besoin que celle-ci soit ajoutée à leur milieu de croissance. Deux autres composés structurellement apparentés à la proline (GSA et la glutamine) ont été ajoutés aux mutants pour tester leur capacité à permettre la croissance des mutants à la place de la proline. Les résultats ont été les suivants : « oui » signifie que le composé a permis la croissance du mutant et « non » qu'il ne l'a pas permise :

Mutant	Proline	Glutamine	GSA
1	Oui	Non	Oui
2	Oui	Non	Non
3	Oui	Oui	Oui
4	Oui	Non	Non
5	Oui	Non	Oui
6	Oui	Non	Oui

- a. Combien de réponses différentes y a-t-il parmi les six mutants ?
- b. Que représentent les différentes réponses ?
- c. Quelle relation existe-t-il entre les trois composés proline, GSA et glutamine ?

6. Le tableau suivant présente des gammes d'activités enzymatiques (en unités dont vous n'avez pas à vous soucier) observées pour des enzymes impliquées dans deux maladies métaboliques récessives chez l'homme. On dispose d'informations similaires pour de nombreuses maladies génétiques métaboliques.

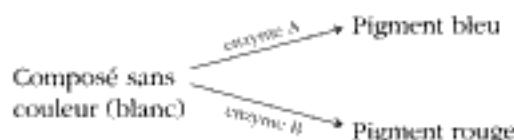
Maladie	Enzyme impliquée	GAMME D'ACTIVITÉ ENZYMATIQUE		
		Malades	Parents des malades	Individus normaux
Acatalasie	Catalase	0	1,2-2,7	4,3-6,2
Galactosémie	Gal-1-P uridyl transférase	0-6	9-30	25-40

- Quelle utilité présente ce type d'information dans le cadre d'un conseil génétique ?
 - Indiquez des sources possibles d'ambiguïté dans l'interprétation de ces études pour un malade isolé.
 - Réexaminez le concept de dominance à la lumière de ces données.
7. Plusieurs mutants de levures ont été isolés. Tous ont besoin du composé G pour croître. Les composés (A à E) dans la voie de biosynthèse qui mène à G sont connus, mais on ignore leur ordre dans cette voie. On teste chaque composé pour voir s'il permet ou non la croissance de chaque mutant (1 à 5). Dans le tableau suivant, + indique la croissance et - l'absence de croissance :
- Quel est l'ordre des composés A à E dans la voie métabolique ?
 - À quel endroit de cette voie chaque mutant est-il bloqué ?

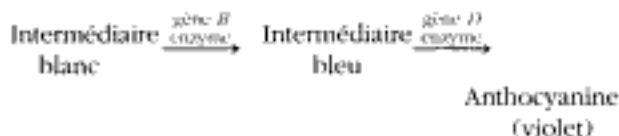
	Composé testé					
	A	B	C	D	E	G
Mutant 1	-	-	-	+	-	+
2	-	+	-	+	-	+
3	-	-	-	-	-	+
4	-	+	+	+	-	+
5	+	+	+	+	-	+

8. Chez une certaine plante, les pétales des fleurs sont violets. Deux mutations récessives apparaissent chez des plantes séparées et se révèlent appartenir à des gènes différents. La mutation 1 (*m1*) donne des pétales bleus chez les homozygotes (*m1/m1*). La mutation 2 (*m2*) donne des pétales rouges chez les homozygotes (*m2/m2*). Les biochimistes travaillant sur la synthèse des pigments des

fleurs de cette espèce ont déjà décrit la voie métabolique suivante :



- À votre avis, chez quel mutant l'activité enzymatique A est-elle déficiente ?
 - Une plante est de génotype *m1'/m1.m2'/m2*. À votre avis, quel est son phénotype ?
 - Pourquoi ces mutants sont-ils récessifs ? (Rappel : le point entre *m1* et *m2'* signifie que l'on considère deux couples d'allèles de deux gènes distincts.)
9. Chez les pois de senteur, la synthèse du pigment anthocyanine violet dans les pétales, est contrôlée par deux gènes, *B* et *D*. La voie de synthèse est la suivante :



- À quelle couleur de pétale vous attendez-vous, pour une plante homozygote présentant une mutation récessive qui l'empêche de catalyser la première réaction ?
 - À quelle couleur de pétale vous attendez-vous, pour une plante homozygote présentant une mutation récessive qui l'empêche de catalyser la deuxième réaction ?
10. Douze mutations complètes d'un gène sans intron de *Neurospora* sont examinées et on découvre que tous les sites mutants sont regroupés dans une région occupant le tiers central du gène. Quelle pourrait être l'explication de cette observation ?
11. Un certain gène de drosophile codant une protéine possède un intron. Si un échantillon important de mutations complètes de ce gène est examiné, y aura-t-il au moins un des sites mutants
- dans les exons ?
 - dans l'intron ?
 - dans le promoteur ?
 - à la limite intron-exon ?

Pour chacun des cas où la réponse est oui, indiquez si le transcrit, le polypeptide, ou les deux, seront décelables.

12. On a obtenu un mutant albinos d'une plante, dépourvu du pigment anthocyanine rouge qui est normalement fabriqué par l'enzyme P. De plus, le tissu de la plante mutante est dépourvu de toute activité détectable de l'enzyme P. Pourtant, une étude immunologique montre clairement la présence d'une protéine présentant les mêmes propriétés immunologiques que l'enzyme P, dans les cellules du mutant. Comment cela est-il possible ?
13. Expliquez les observations suivantes faites sur trois gènes de drosophile. Dans tous les cas, inventez des symboles pour les allèles.
- On trouve normalement vingt unités de l'enzyme E dans les cellules homozygotes de type sauvage. Un mutant homozygote présente zéro unité. Un hétérozygote a un aspect mutant et possède 10 unités de E par cellule.
 - On trouve normalement trente unités de l'enzyme F dans les cellules homozygotes de type sauvage. Un mutant homozygote présente zéro unité. Un hétérozygote a un aspect sauvage et possède 15 unités de E par cellule.
14. Chez un organisme diploïde imaginaire, la forme carrée des cellules est due à un effet de seuil : plus de 50 unités du « facteur carré » par cellule produiront un phénotype carré, et moins de 50 unités produiront un phénotype rond. L'allèle *sf* (*sf* : *square factor* : facteur carré) est un gène responsable de la synthèse du facteur carré. Chaque allèle *sf* contribue pour 40 unités du facteur carré. Donc, les homozygotes *sf/sf* présentent 80 unités et ont un phénotype carré. Un allèle mutant *sn* (*sn* : *square null* : pas de carré) apparaît ; il n'est pas fonctionnel et n'apporte aucun facteur carré.
- Quel allèle présentera une dominance, *sf* ou *sn* ?
 - Les allèles fonctionnels sont-ils nécessairement toujours dominants ? Justifiez votre réponse.
 - Dans un système tel que celui-ci, comment un allèle particulier peut-il subir un changement au cours de l'évolution de sorte que son phénotype présente une transmission récessive à la génération 0 et une transmission dominante à la génération suivante ?
15. En Norvège en 1934, une mère ayant deux enfants présentant un retard mental a consulté le médecin Asbjørn Folling. Au cours de l'entretien, Folling apprit que l'urine des enfants avait une odeur curieuse. Il testa cette urine avec du chlorure ferrique et découvrit que l'urine des enfants restait verte au lieu de devenir marron comme l'urine normale. Il déduisit de ceci que le responsable chimique devait être l'acide phénylpyruvique. En raison de sa ressemblance chimique avec la phénylalanine, il semblait probable que cette substance se soit formée à partir de la phénylalanine présente dans le sang, mais il n'existait à ce moment-là aucun moyen de mettre la phénylalanine en évidence. Toutefois, une bactérie pouvait convertir la phénylalanine en acide phénylpyruvique. On pouvait donc mesurer le taux de phénylalanine en utilisant le test au chlorure ferrique. On découvrit que les enfants présentaient une concentration élevée de phénylalanine dans le sang, ce qui était probablement à l'origine de l'acide phénylpyruvique. À partir d'études menées sur des familles similaires, Folling déduisit que cette maladie, que l'on appelle maintenant phénylcétonurie (PCU), semblait avoir une origine génétique et se présentait comme un caractère récessif simple. Il devint évident que la phénylalanine était la coupable, que cette substance chimique s'accumulait chez les malades atteints de PCU et était transformée, atteignant des concentrations élevées en acide phénylpyruvique. Ceci interférait alors avec le développement normal du tissu nerveux. Cette observation conduisit à la mise au point d'un régime alimentaire spécial, pauvre en phénylalanine, que l'on pouvait donner aux nouveau-nés chez lesquels on avait diagnostiqué une PCU et qui permettait leur développement sans retard psychomoteur. On découvrit également qu'après le développement du système nerveux chez l'enfant, le malade pouvait abandonner son régime spécial. Malheureusement, de nombreuses femmes atteintes de PCU qui s'étaient développées normalement grâce au régime spécial eurent des enfants présentant un retard mental dès la naissance, et sur lesquels le régime spécial mis au point n'avait aucun effet.
- À votre avis, pourquoi les bébés des mères atteintes de PCU présentaient-ils un retard mental à la naissance ?
 - Pourquoi un régime spécial n'avait-il aucun effet sur eux ?
 - Expliquez la raison de la différence de résultats entre les bébés PCU et les bébés des mères PCU.
 - Proposez un traitement qui pourrait permettre aux mères PCU d'avoir des enfants normaux.
 - Écrivez un texte court sur la PCU, en y incorporant les différents concepts aux niveaux génétique, diagnostique, enzymatique et physiologique.
16. L'addition d'un nucléotide et la délétion d'un nucléotide environ 15 sites plus loin dans l'ADN modifie la séquence de la protéine suivante :
- Lys-Ser-Pro-Ser-Leu-Asn-Ala-Ala-Lys
- en
- Lys-Val-His-His-Leu-Met-Ala-Ala-Lys

- Quelles sont l'ancienne et la nouvelle séquences nucléotidiques de l'ARNm ? (Utilisez le dictionnaire des codons).
- Quel nucléotide a été ajouté et quel nucléotide a disparu ?

(Problème 16 d'après W. D. Stansfield. *Theory and problems of genetics*, McGraw-Hill, 1969.)

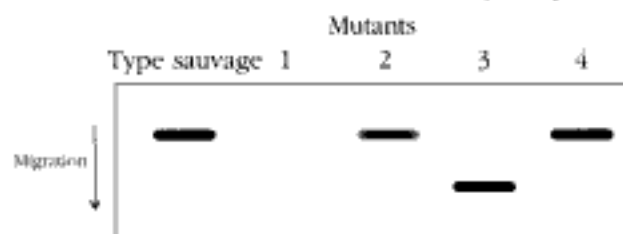
- Normalement la thyroxine, une hormone de croissance thyroïdienne, est fabriquée dans le corps par une enzyme, de la façon suivante :



Si l'enzyme est défectueuse, elle entraîne des symptômes qui constituent le syndrome crétino-goitreux (GGC pour *genetic goiterous cretinism*), un syndrome rare qui consiste en une croissance lente, une thyroïde surdimensionnée (goitre) et un retard mental.

- Si l'allèle normal est haplo-suffisant, pensez-vous que le GGC sera transmis sous la forme d'un phénotype dominant ou récessif ? Expliquez pourquoi.
 - Faites des hypothèses sur la nature de l'allèle responsable du GGC, en comparant sa séquence moléculaire à celle de l'allèle normal. Dans le cadre de votre modèle, montrez pourquoi cela induit une enzyme inactive.
 - Comment les symptômes du GGC pourraient-ils être atténués ?
 - À la naissance, les enfants atteints de GGC sont parfaitement normaux et ne développent leurs symptômes que plus tard. À votre avis, pourquoi ?
- Indiquez les points communs et les différences qui existent entre les processus au cours desquels l'information devient forme, dans un organisme et lors de la construction d'une maison.
 - Essayez de trouver des exceptions à l'affirmation du chapitre 1 où il est écrit que « lorsque vous regardez un organisme, ce que vous voyez est soit une protéine, soit quelque chose qui a été fabriqué par une protéine ».
 - On a obtenu quatre mutants chez la levure, tous dépourvus de l'activité d'une enzyme spécifique E dont la séquence polypeptidique primaire est codée par un seul gène. On a montré que toutes les mutations étaient contenues dans ce gène. L'électrophorèse est une technique au cours de laquelle les macromolécules telles que l'ADN, l'ARN et les protéines sont séparées selon leur taille et leur charge dans un milieu gélatineux (un « gel ») en établissant un champ électrique élevé dans celui-ci.

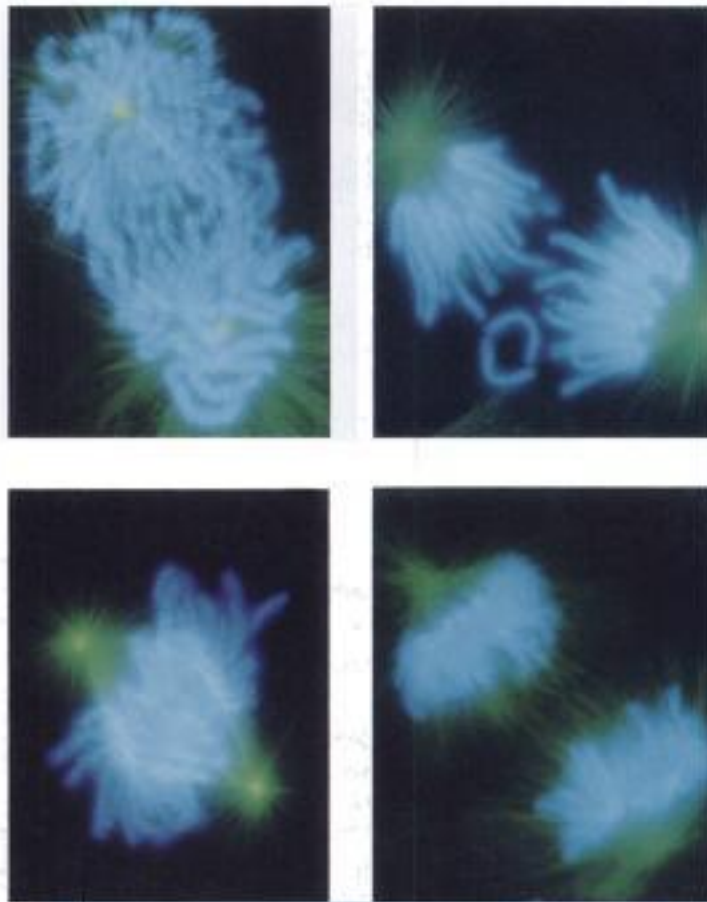
La protéine a été extraite des mutants et d'un témoin de type sauvage. Après la séparation par électrophorèse, le gel a été marqué par un anticorps couplé à un colorant, spécifique de l'enzyme E. Les résultats sont présentés dans la colonne ci-contre. Donnez une explication moléculaire des résultats de chaque piste (colonne de migration) et montrez de quelle façon chacune est compatible avec l'absence observée d'activité enzymatique.



- Inventez des symboles pour les mutations géniques suivantes et pour leurs allèles de type sauvage (et expliquez le choix de ces symboles) :
 - Une mutation auxotrophe chez *Neurospora* pour l'acide aminé histidine (le type sauvage est prototrophe pour l'histidine).
 - Une mutation récessive chez la drosophile entraînant un corps jaune (le type sauvage a un corps gris).
 - Une mutation dominante chez le maïs qui produit un pigment violet (le type sauvage n'en produit pas).
 - Une mutation complète d'un gène haplo-insuffisant qui spécifie normalement les poils sur la tige des pétunias.
 - Indiquez le génotype d'un hétérozygote dans les cas b à d.
- Citez trois exemples de complémentarité des bases dans la fonction des gènes.
- Citez trois exemples de protéines agissant sur les acides nucléiques.
- Un plasmide linéaire contient deux gènes non chevauchants transcrits en sens inverse, chacun à partir d'une extrémité et vers le centre du plasmide. Dessinez des schémas qui montrent :
 - L'ADN plasmidique, en indiquant les extrémités 5' et 3' des brins nucléotidiques.
 - Le brin matrice pour chaque gène.
 - Les positions des sites d'initiation de la transcription.
 - Les transcrits, en indiquant leurs extrémités 5' et 3'.
 - Les positions des codons d'initiation et des codons stop.
 - Les chaînes d'acides aminés correspondantes, alignées avec les transcrits et l'identification des extrémités carboxyle et aminée.

4

LA TRANSMISSION DES GÈNES



POUR L'ESSENTIEL

- La réplication de l'ADN précède la division cellulaire sexuée et asexuée.
- L'ADN se réplique en se déroulant et les brins exposés sont utilisés comme matrices pour la nouvelle synthèse. Chaque molécule est donc formée d'un brin d'origine et d'un brin néosynthétisé. L'ADN polymérase ajoute des nucléotides aux extrémités 3' en cours d'élongation.
- Dans le cas de la division cellulaire asexuée, la division nucléaire concomitante (mitose) produit deux cellules filles génétiquement identiques.
- Les cellules sexuelles haploïdes sont produites par deux divisions consécutives d'un méiocyte diploïde, qui s'accompagnent de deux divisions nucléaires (méiose).
- Les rapports phénotypiques dans la descendance sont déterminés par les mouvements des paires de chromosomes lors de la méiose.
- Les membres d'une paire d'allèles hétérozygotes subissent une ségrégation égale dans les cellules sexuelles.
- Les gènes situés sur le chromosome X peuvent présenter des rapports phénotypiques différents dans la descendance mâle et femelle.
- Les gènes des organites sont transmis par la mère.

Chromosomes de mammifères à différents stades de la mitose. La couleur bleue vient d'un colorant spécifique de l'ADN, l'homodimère d'éthidium et la couleur verte d'un anticorps, couplé à un fluorochrome, qui se fixe à la tubuline, un composant protéique des fibres du fuseau.

(D'après J. C. Waters, R. W. Cole et C. L. Rieder, 1993, *J. Cell Biol.* 122 : 361, merci à C. L. Rieder.)

L'un des plus grands mystères de l'hérédité à travers les âges a été la transmission des caractères. Il y a deux grandes interrogations dans ce mystère. Premièrement, la conservation des espèces : les hommes enfantent toujours des bébés et les chats des chatons. Deuxièmement, la transmission de phénotypes différents à travers une même espèce : nous en connaissons tous des exemples tels que « elle a hérité du nez de son père » ou « il y a des cheveux roux dans la famille ». Quel est le mécanisme par lequel les caractères et les formes des caractères (phénotypes) sont transmis ? Dans ce chapitre, nous étudierons les mécanismes de transmission au niveau des individus, des cellules, des chromosomes et des gènes. À l'origine de la transmission à tous ces niveaux se trouve la réplication de l'ADN. La réplication est le processus au cours duquel des répliques ou copies sont fabriquées. Une molécule d'ADN peut se répliquer en deux molécules « filles » identiques, et c'est de cette façon que les chromosomes peuvent se dupliquer, qu'une cellule peut se diviser en deux, qu'un nouvel organisme pluricellulaire peut être formé, de cette façon aussi que les populations s'agrandissent et que les organismes se perpétuent. L'ADN que nous portons est l'aboutissement d'une série de réplifications, ininterrompue depuis des milliards d'années.

LA RÉPLICATION DE L'ADN

Tant chez les Procaryotes que chez les Eucaryotes, la réplication de l'ADN est un préalable à la division cellulaire. Cette phase de réplication de l'ADN est appelée **phase S (de synthèse)**. Les deux molécules filles d'ADN qui se sont formées au cours de la réplication deviennent des chromosomes à part entière dans les cellules filles.

Comme pour tous les phénomènes qui impliquent les acides nucléiques, la machinerie élémentaire de réplication de l'ADN dépend de la complémentarité des molécules d'ADN et de la capacité de certaines protéines à établir des interactions spécifiques avec des séquences particulières de l'ADN.

La réplication semi-conservative

La figure 4-1 schématise le mécanisme de réplication de l'ADN chez les Procaryotes et les Eucaryotes. Imaginez la double hélice comme une fermeture éclair qui s'ouvre en partant d'une extrémité (le bas de cette figure). On peut voir que le déroulement des deux brins expose des bases non appariées sur chaque brin. Ainsi, chacune des bases exposées peut s'apparier avec les nucléotides libres présents en solution. (Ces nucléotides ajoutés proviennent d'une réserve synthétisée chimiquement dans le cytoplasme.) Les règles d'appariement des bases sont strictes, c'est pourquoi chacune des bases exposées peut s'apparier uniquement avec la base qui lui est complémentaire. En raison de cette complémentarité des bases, chacun des

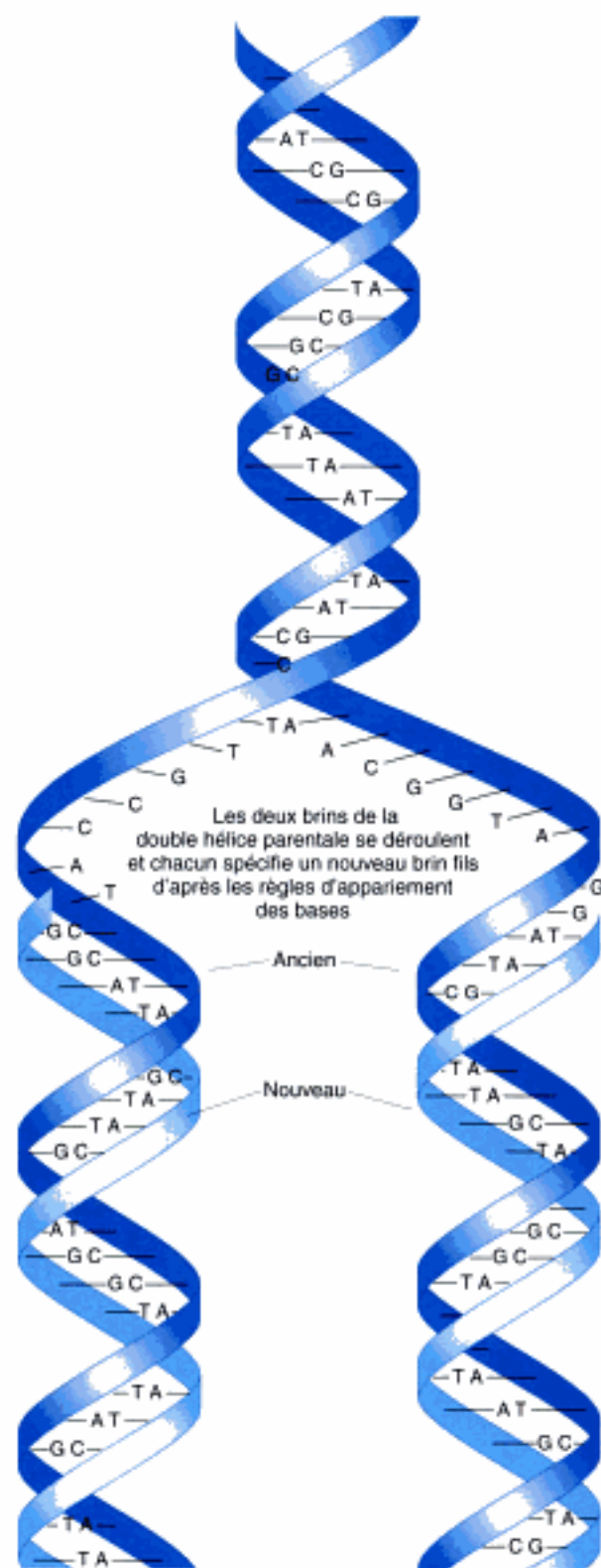


Figure 4-1 Le modèle de réplication de l'ADN proposé par Watson et Crick est basé sur la spécificité de formation des liaisons hydrogène entre les paires de bases. Les brins complémentaires sont représentés dans des couleurs différentes.

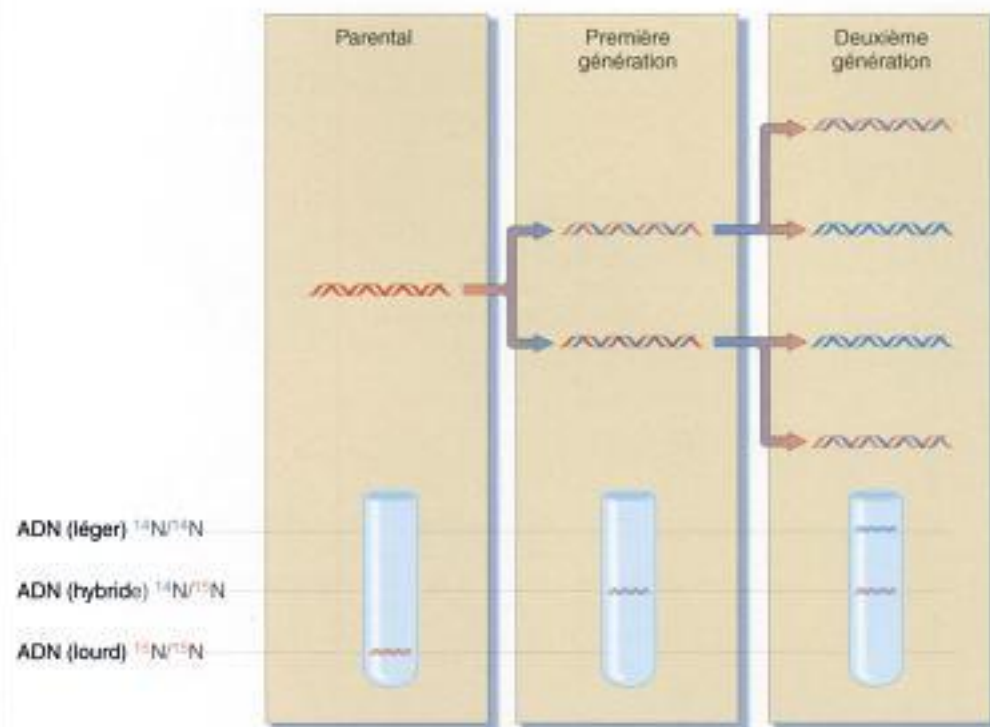
deux brins séparés sert de **matrice** (guide d'alignement) pour reformer une double hélice identique à l'originale. On dit que cette réplication est **semi-conservative** car chaque double hélice fille comporte un brin parental et un brin néo-synthétisé. Si l'on autorise des molécules d'ADN à se répliquer parmi des nucléotides contenant des isotopes rares (qui servent de marqueurs), on peut montrer expérimentalement que les molécules filles contiennent un brin avec l'isotope normal et un brin avec la forme rare (voir Les grandes découvertes de la génétique 4-1).

de matrices pour la polymérisation de nouveaux brins, conduisant ainsi à la formation de deux double hélices filles. Ce type de réplication est devenu célèbre sous le nom de **réplication semi-conservative**. Une belle idée, mais de nombreuses idées séduisantes ont été mises à mal par des faits réels. Il était donc nécessaire de vérifier cette suggestion. Le premier élément de preuve fut apporté par Matthew Meselson et Frank Stahl.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 4-1

Meselson et Stahl montrent que l'ADN se réplique de façon semi-conservative

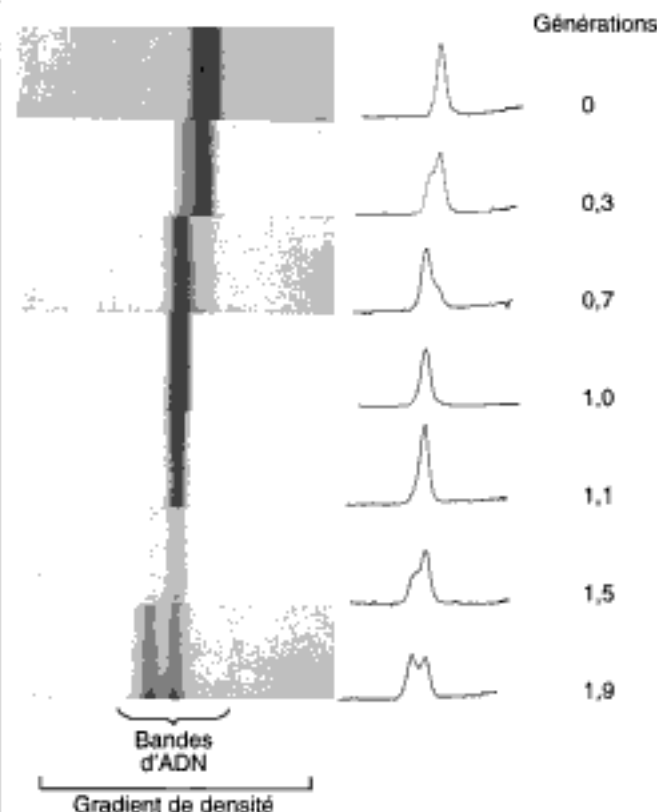
Dans leur article de 1953 décrivant pour la première fois la structure de l'ADN, Watson et Crick terminèrent par le bref commentaire énigmatique : « Il n'a pas échappé à notre attention que l'appariement spécifique dont nous avons fait l'hypothèse suggère immédiatement un mécanisme possible de copie du matériel génétique. » Pour les généticiens de cette époque, la signification de ce communiqué a dû être claire : les deux moitiés de la double hélice doivent se séparer et, en raison de la spécificité de l'appariement des bases, elles doivent toutes deux servir



Meselson et Stahl écrivirent plus tard :

Nous nous sommes rencontrés lorsque nous étions étudiants en troisième cycle, à Woods Hole, en 1954, le deuxième été suivant l'annonce du modèle de Watson et Crick. Nos premières conversations portèrent sur les solutions de certaines intégrales décrivant la réactivation croisée de phages induits par des UV, que l'un de nous étudiait alors. Plus tard, nous en vîmes à parler des possibilités de vérifier l'hypothèse de la duplication semi-conservative de l'ADN. Les idées les plus farfelues ont sans doute été oubliées, mais nous nous rappelons avoir parlé de l'utilisation possible de la densité comme marquage des molécules d'ADN parental.

L'idée consistait à permettre à des molécules parentales contenant des composants d'une certaine densité, de se répliquer dans un milieu contenant des composants de densités



différentes. Si l'ADN se répliquait de façon semi-conservative, les molécules filles devaient avoir une densité intermédiaire. Les premières tentatives de cette approche avec les phages, en utilisant l'isotope lourd de l'hydrogène, le deutérium et l'analogue de la thymine, le 5-bromo-uracile (5BU) échouèrent, mais

Nous avons finalement abandonné le phage et le 5BU et nous nous sommes tournés vers les bactéries et l'isotope lourd de l'azote (^{15}N). Les deuxième et troisième expériences de ce type fonctionnèrent magnifiquement. Nous les avons donc renommées 1 et 2 et nous avons commencé à écrire l'article.

L'article fut publié en 1958.

Des densités différentes d'ADN peuvent être visualisées grâce aux positions que l'ADN prend dans un gradient de densité de chlorure de césium, formé lors d'une ultracentrifugation. (Les forces gravitationnelles élevées créées dans l'ultracentrifugeuse par une rotation à haute vitesse sont nécessaires pour établir le gradient de densité.) Les cellules cultivées initialement dans un milieu contenant du ^{15}N présentent de l'ADN de densité élevée. Après avoir cultivé ces cellules dans du ^{14}N pendant une génération, les chercheurs découvrirent que l'ADN était de densité intermédiaire et qu'après deux générations, on observait de l'ADN de densité intermédiaire et de l'ADN de faible densité, ce qui correspondait exactement à ce que prédisait le modèle de Watson et Crick.

Le processus de polymérisation

Commençons par une vue d'ensemble des événements de polymérisation qui se produisent au cours du processus de réplication. L'enzyme qui catalyse la polymérisation des nucléotides s'appelle l'**ADN polymérase**. Cette enzyme ajoute des désoxynucléotides à l'extrémité 3' d'une chaîne nucléotidique en cours d'élongation, en utilisant une matrice d'ADN simple-brin (figure 4-2). Nous avons vu au chapitre 3 que l'ARN polymérase agissait d'une manière similaire, en ajoutant des ribonucléotides à l'extrémité 3' d'une chaîne d'ARN en cours d'élongation. Les substrats de l'ADN polymérase sont les formes triphosphatées des désoxynucléotides dATP, dGTP, dCTP et dTTP. L'ADN polymérase intervient au niveau de la **fourche de réplication**, la zone au niveau de laquelle l'ADN se déroule et expose ses simples-brins pour qu'ils servent de matrices. La polymérisation des nucléotides catalysée par l'ADN polymérase ayant toujours lieu au niveau de l'extrémité 3', la nouvelle synthèse ne peut avoir lieu de façon continue que sur l'un des brins matrice. Le nouveau brin ainsi formé est appelé **brin précoce** (figure 4-3). La synthèse sur l'autre matrice se fait également à partir de l'extrémité 3' mais elle procède par courts fragments et se déroule « dans le mauvais sens » (car pour ce brin, le sens 5'→3' va en s'éloignant de la fourche de réplication). Pour ce brin, l'ADN polymérase s'éloigne de la matrice exposée et doit attendre que la

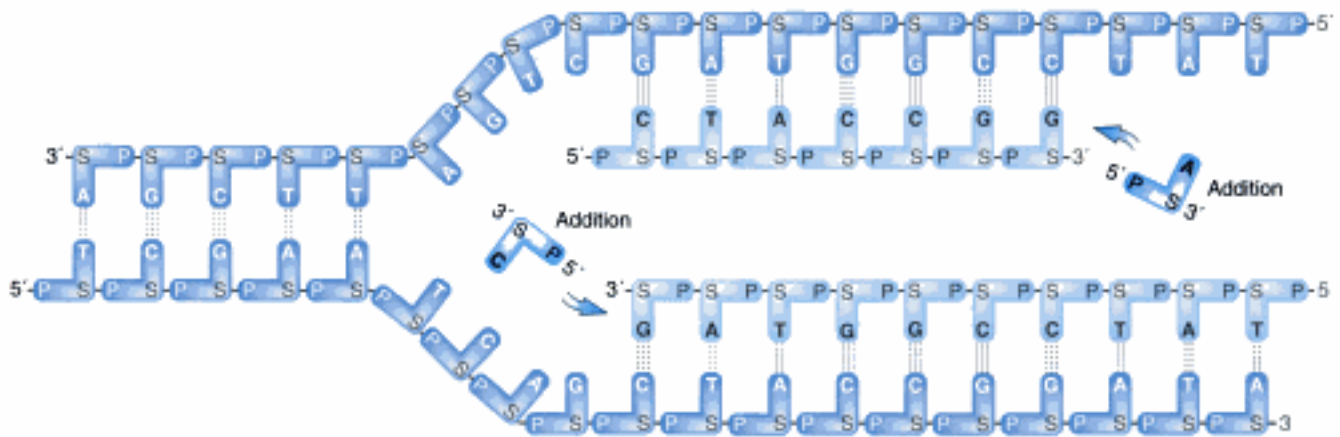


Figure 4-2 Fourche de réplication de l'ADN montrant de façon simplifiée la manière dont les nucléotides sont ajoutés aux extrémités 3' en cours d'élongation, sur chaque matrice mais en sens opposés. Voir la figure 4-3 pour plus de détails et pour les étapes suivantes.

fourche de réplication expose davantage d'ADN avant de pouvoir synthétiser un autre court fragment d'ADN. Ces fragments, appelés **fragments d'Okazaki** sont ensuite reliés les uns aux autres par une enzyme appelée **ADN ligase**. Le nouveau brin ainsi formé est le **brin retardé**.

De nombreuses protéines actives sont nécessaires pour exécuter le processus général de réplication décrit ci-dessus. Les composants impliqués dans la réplication chez la bactérie *Escherichia coli* sont présentés dans la figure 4-4 à la page suivante. La principale polymérase est l'ADN polymérase III (pol III), qui catalyse l'addition des nucléotides en 3' au niveau de la fourche de réplication. Toutefois, les polymérases de ce type ne peuvent ajouter des nucléotides qu'à des chaînes nucléotidiques préexistantes. C'est pourquoi au début de la réplication des brins précoce et retardé, de courts fragments d'ARN sont synthétisés pour servir de points de départ, d'**amorces** à la polymérisation. Pour le brin précoce, une seule amorce initiale est nécessaire car après le premier amorçage, l'addition continue peut utiliser le brin d'ADN en cours de synthèse comme amorce. En revanche sur le brin retardé, chaque fragment d'Okazaki a besoin de sa propre amorce. Les amorces sont synthétisées par un ensemble de protéines appelé un **primosome**, dont un des composants principaux est une enzyme, la **primase**, un type

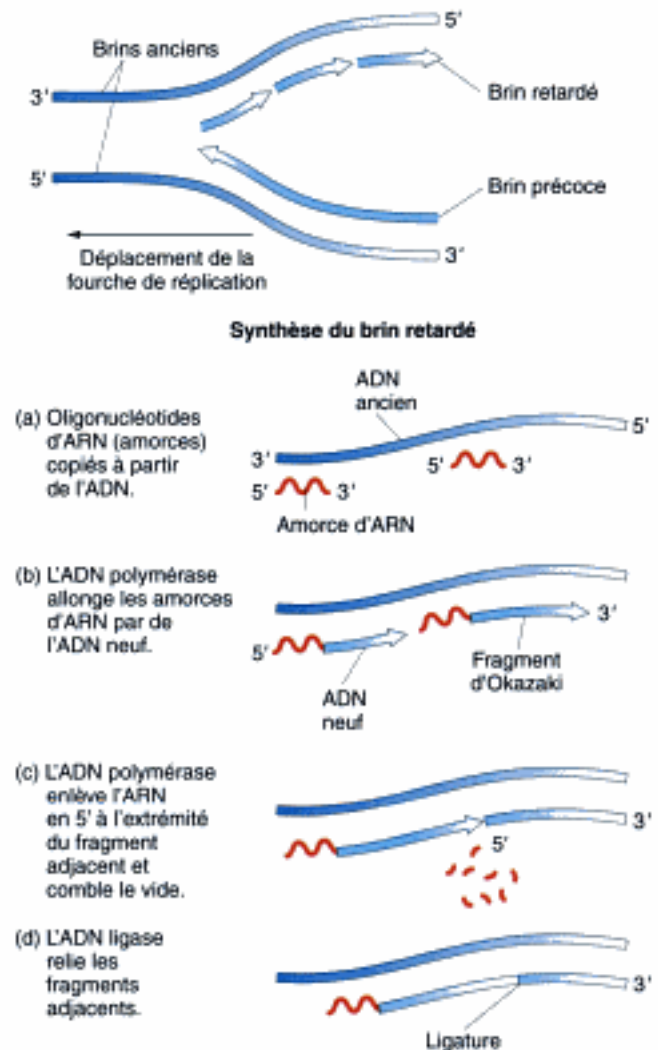


Figure 4-3 La structure complète d'une fourche de réplication (en haut) et les étapes de la synthèse du brin retardé. (D'après H. Lodish, D. Baltimore, A. Berk, S. L. Zipursky, P. Matsudaira et J. Darnell, *Biologie moléculaire de la cellule*. Traduction française de la 3^e édition chez De Boeck, 1997.)

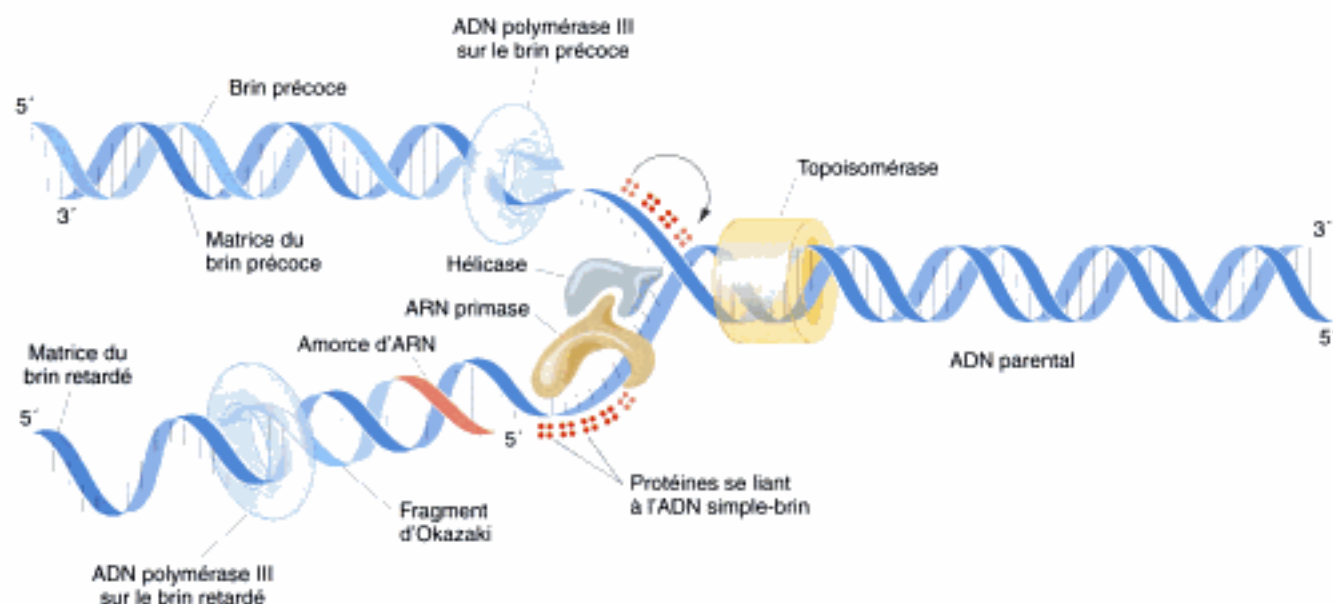


Figure 4-4 Fourche de réplication de l'ADN. (D'après Purves, Orians, Heller. *Life : The science of biology*, 4^e éd. Copyright © 1995 par Sinauer Associates, Inc.).

d'ARN polymérase. Le retrait des amorces d'ARN et le remplissage par de l'ADN, des vides laissés par ce retrait est assuré par une autre ADN polymérase, pol I. Après que pol I a rempli son rôle, une ligase relie l'extrémité 3' de l'ADN, comblant les vides à l'extrémité 5' du fragment d'Okazaki situé en aval.

Le déplacement de la fourche de réplication est assuré par une enzyme, l'**hélicase**, qui rompt les liaisons hydrogène établies entre les bases appariées et déroule la double hélice devant l'ADN polymérase en cours de progression. Des protéines capables de se lier à l'ADN simple-brin empêchent les simples-brins d'ADN ainsi créés de se réassocier. Lorsque l'ADN se déroule, il a tendance à former des supertours, un processus similaire à celui que l'on observe lorsqu'on essaie de séparer deux brins d'une corde. La double hélice est ramenée à son état de départ (sans supertour) par l'action d'une autre enzyme, la **gyrase**, qui est un type de **topoisomérase**. Cette classe d'enzymes peut couper et relier des brins d'ADN, leur permettant de « passer au travers » l'un de l'autre, comme un magicien qui entrelace et sépare des anneaux d'acier.

Les origines de réplication

Dans les génomes des Procaryotes et des Eucaryotes, la réplication débute au niveau de séquences nucléotidiques spécifiques reconnues par la machinerie de réplication. On les appelle des **origines de réplication**. La synthèse se déroule alors de façon bidirectionnelle : deux fourches avancent en sens inverse, comme le montre la figure 4-5a. Les double hélices répliquées produites par chaque origine de réplication s'allongent et finissent par se rejoindre. Lorsque la réplication des deux brins est terminée, on obtient deux **molécules filles** identiques d'ADN. Dans le cas d'un chromosome eucaryote, ces répliques sont appelées **chromatides sœurs** (figure 4-5b). Notez que le terme *chromatide* ne s'applique que transitoirement. Les chromatides sont en fait d'authentiques chromosomes et elles reprennent leur identité après la division cellulaire.

MESSAGE

L'ADN se réplique de façon semi-conservative. Grâce à l'utilisation comme matrices, des simples-brins séparés de la double hélice, les nucléotides sont polymérisés au niveau des extrémités 3' des nouvelles chaînes. L'addition se déroule de façon continue pour le brin précoce mais elle a lieu de façon discontinue pour le brin retardé.

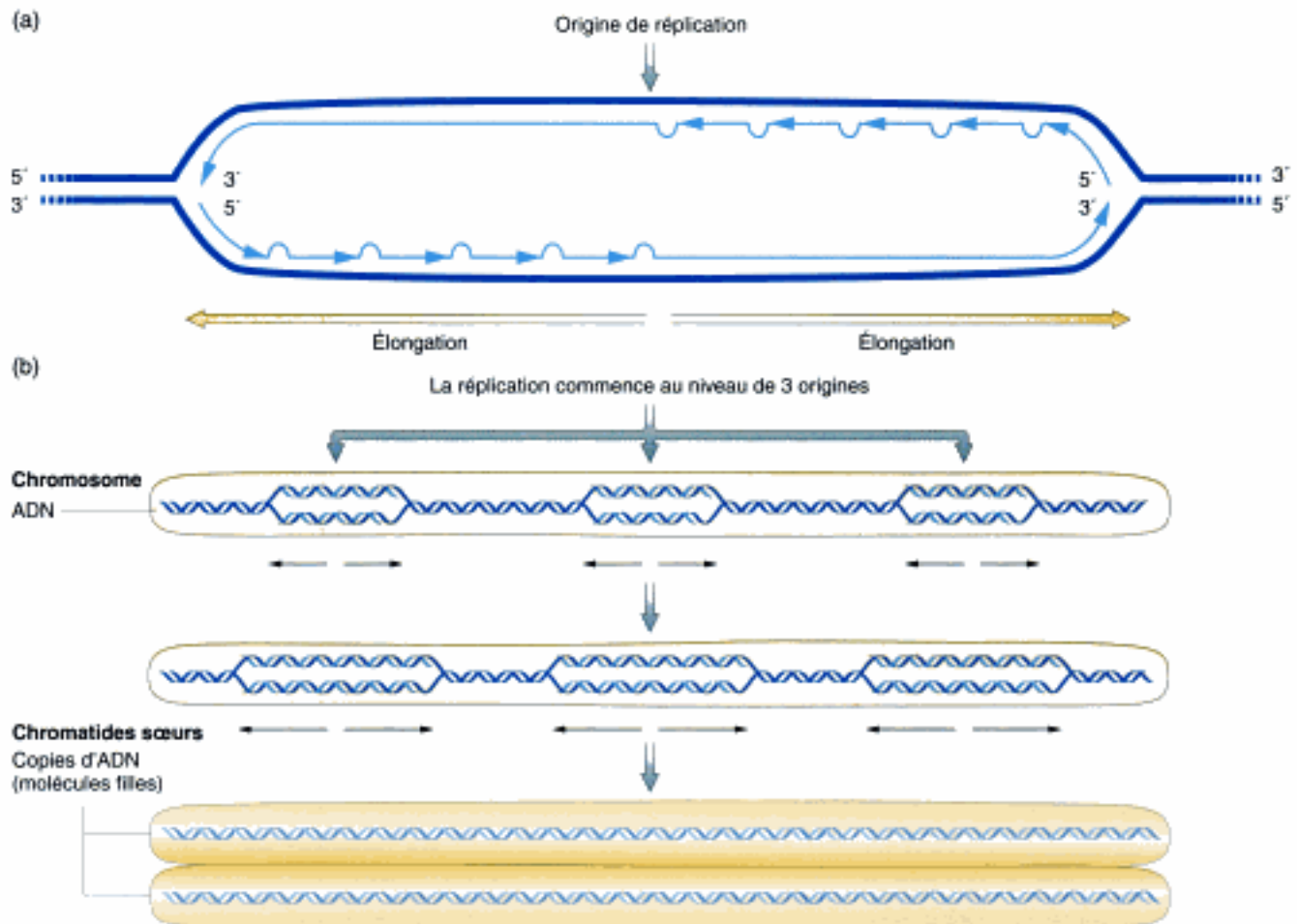


Figure 4-5 (a) La nature bidirectionnelle de la réplication de l'ADN. Commençant au niveau de l'origine de réplication, les ADN polymérases s'en éloignent dans les deux sens. Les flèches longues montrent les brins précoces et les flèches courtes qui se touchent indiquent les brins retardés. (b) Déroulement de la réplication au niveau du chromosome. Trois origines sont présentées dans cet exemple).

LA DIVISION CELLULAIRE

La division cellulaire est à la base de toutes les formes de reproduction des organismes. Les organismes unicellulaires se divisent pour se reproduire. La division cellulaire chez les organismes pluricellulaires produit des cellules reproductrices spécialisées, telles que les ovules et les spermatozoïdes et est également responsable du développement d'un organisme pluricellulaire à partir d'un œuf fécondé unique. Pour que la cellule se divise, le génome doit également se diviser. C'est pourquoi dans tous les types de divisions cellulaires chez tous les organismes, la réplication de l'ADN précède la division cellulaire.

Les principaux types de division cellulaire sont présentés dans la figure 4-6 à la page suivante. On peut grossièrement les regrouper en division cellulaire sexuée ou asexuée.

La division cellulaire asexuée

Chez les Procaryotes, il existe un seul type de division cellulaire assez simple, qui produit deux cellules filles identiques à partir d'une cellule mère. On parle de **division cellulaire asexuée** car elle ne fait pas intervenir d'union sexuelle entre des individus différents. Avant la division cellulaire bactérienne, l'ADN se réplique et forme deux génomes circulaires complets. Lorsque la cellule se divise, l'une des molécules d'ADN filles passe dans chacune des cellules filles. Le mécanisme précis de ce déplacement de l'ADN est inconnu, mais dans un modèle, l'ADN est ancré à la membrane cellulaire et lorsque la membrane s'allonge pour fabriquer deux cellules, l'ADN est entraîné à sa suite.

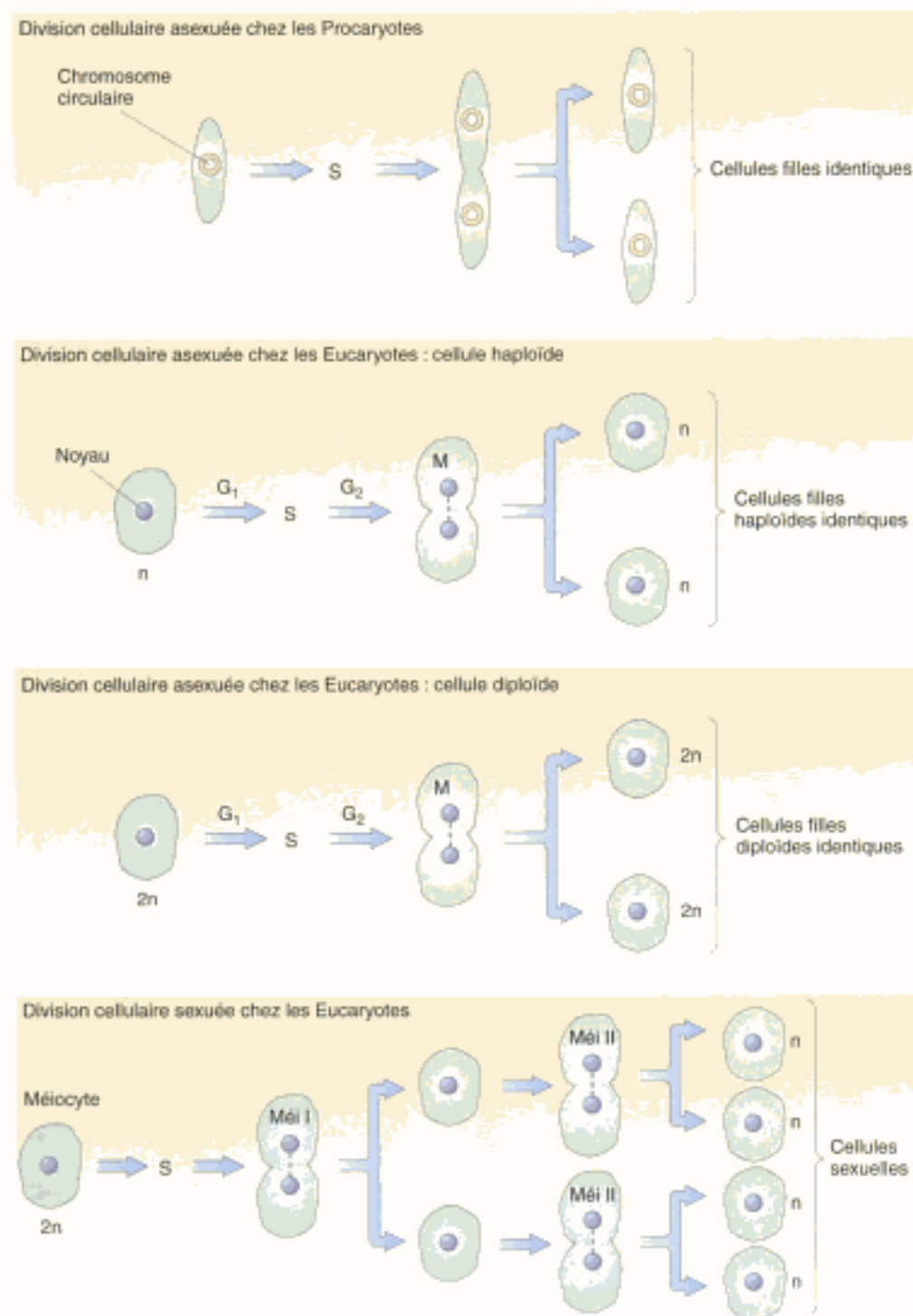


Figure 4-6 Les différents types de divisions cellulaires et les divisions des génomes qui leurs sont associées. M = mitose ; Méi = méiose.

On observe également des divisions cellulaires asexuées chez les Eucaryotes, qui conduisent aussi à deux cellules filles identiques à partir d'une cellule mère. C'est le type de division cellulaire qui convertit un œuf fécondé (une seule cellule), le **zygote**, en deux cellules, puis en quatre, en huit et ainsi de suite jusqu'à ce qu'un orga-

nisme pluricellulaire soit produit. C'est également le type de division cellulaire qui produit une population d'organismes unicellulaires à partir d'un seul, comme par exemple, chez les levures et les protozoaires. Lorsqu'une cellule eucaryote se divise de façon asexuée, le noyau et son contenu génétique doivent également se diviser, par

un processus appelé **mitose**. Les cellules, tant haploïdes (n) que diploïdes ($2n$) peuvent se diviser de façon asexuée, auquel cas leur noyau subit une division mitotique. La figure 4-7 indique les moments où les divisions cellulaires asexuée (mitotique) et sexuée (méiotique) se produisent au cours des cycles vitaux de l'homme, des végétaux et des champignons. Lors de la mitose, les déplacements programmés des chromosomes garantissent que chaque cellule fille reçoit des chromosomes et des gènes identiques à ceux de l'autre cellule fille et de la cellule mère.

Le temps et les événements qui séparent n'importe quel stade spécifique chez une cellule mère, du même stade chez une cellule fille est appelé **cycle de division cellulaire**. Il est constitué de quatre stades : **S** (synthèse d'ADN), **M** (mitose) et **G1** et **G2** (stades intermédiaires, gap en anglais) comme on l'a vu dans la figure 4-6. Le passage de la cellule à travers chacun de ces stades successifs est un processus régulé de façon précise, contrôlé par une multitude de gènes divers dont la fonction est d'assurer le déroulement correct de cette séquence.

Au cours de la phase S, l'ADN de chaque chromo-

some est répliqué de façon semi-conservative. Le chromosome se sépare longitudinalement en deux pour produire une paire de chromatides sœurs, dont chacune contient l'une des molécules répliquées d'ADN. Les chromatides deviennent visibles au microscope seulement au moment de la mitose, où elles se raccourcissent et s'épaississent à la suite d'un enroulement plus important de l'ADN et des histones qui lui sont associées. Cependant, il faut se rappeler que la réplication en chromatides a eu lieu au cours de la phase prémitotique S. Les chromatides et l'ADN répliqué sont représentés dans la figure 4-8 à la page suivante. La figure présente les exemples de cellules diploïdes de génotype b'/b' , b'/b et b/b et de cellules haploïdes de génotype b' et b . On suppose que l'allèle mutant b a été formé à la suite du remplacement d'une paire de bases G•C dans l'allèle de type sauvage, par une paire A•T dans l'allèle mutant (c'est un type de mutation courant).

Au cours de la mitose, un réseau de microtubules protéiniques appelé le *fuseau*, se forme parallèlement à l'axe de la cellule, reliant ainsi les deux pôles de la cellule. Un ou plusieurs microtubules émis par un pôle

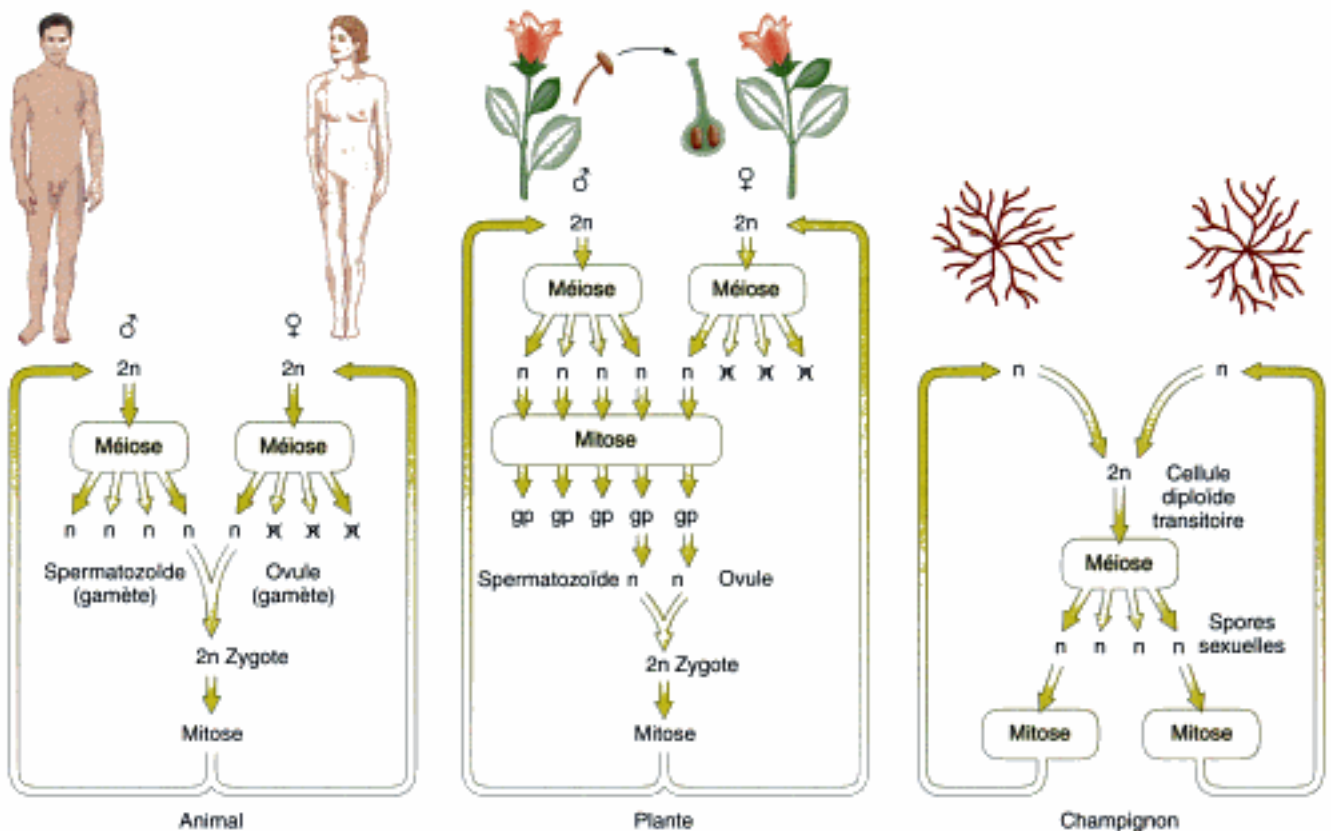


Figure 4-7 Cycles vitaux de l'homme, des végétaux et des champignons, montrant les divisions au cours desquelles se produisent la mitose et la méiose. Remarquez que chez l'homme et chez de nombreux végétaux, trois cellules de la tétrade méiotique avortent. L'abréviation n indique une cellule haploïde, $2n$, une cellule diploïde. gp désigne un gaméophyte, la petite structure de cellules haploïdes qui produira les gamètes. Chez les plantes à fleurs, le stade gaméophyte est radicalement réduit mais chez les autres plantes (telles que les mousses) le gaméophyte est le principal stade végétatif.

Réplication de l'ADN pour former des chromatides

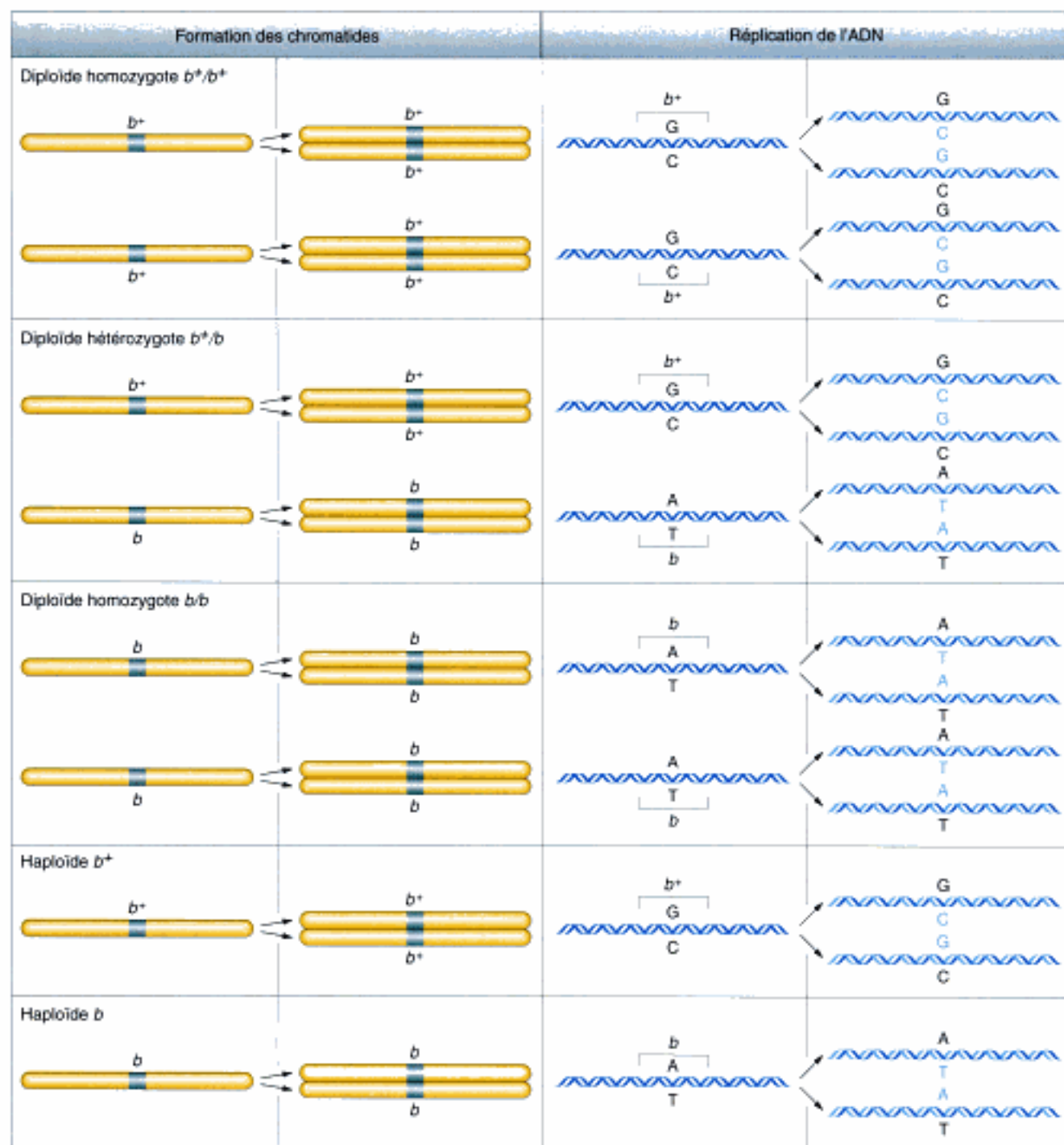


Figure 4-8 Formation des chromatides et réplication de l'ADN sous-jacente. (À gauche) Chaque chromosome se divise longitudinalement en deux chromatides ; (à droite) au niveau moléculaire, l'unique molécule d'ADN dont est constitué chaque chromosome se réplique, produisant deux molécules d'ADN, une pour chaque chromatide. On voit également un allèle b^+ et sa forme mutante b , créée par un changement d'une seule paire de bases G•C en A•T.

s'attachent à une chromatide et un nombre similaire de microtubules émis par le pôle opposé se fixent à l'autre chromatide du chromosome. Le point de fixation sur le chromosome est le centromère, une séquence spécifique d'ADN, essentielle au déplacement des chromatides pendant la mitose. Le centromère est répliqué au cours de la formation des chromatides sœurs et chaque centromère frère sert de site de liaison au complexe multiprotéique appelé le **kinétochore**. Les kinétochores agissent à leur tour comme des sites de fixation pour les microtubules. Les fibres du fuseau (microtubules) tirent ensuite les chromatides sœurs vers les pôles opposés*. Chaque pôle reçoit donc une copie de chaque chromosome en provenance de la cellule mère. Les jeux de chromatides sœurs présents à chaque pôle sont incorporés dans le noyau de chaque cellule fille. Ces noyaux fils sont identiques l'un à l'autre et au noyau dont ils sont issus. Dans les cellules filles, les chromatides sont à nouveau appelées **chromosomes**. Remarquez que c'est le fuseau et le complexe centromère-kinétochore qui déterminent la fidélité de la division nucléaire.

Les principaux événements génétiques de la phase S et de la mitose sont schématisés dans la colonne de gauche et la colonne centrale de la figure 4-9, à la page 97. La figure présente un haploïde de génotype *A* et un diploïde de génotype *A/a* pour montrer que non seulement la mitose produit deux cellules chromosomiquement identiques mais aussi que l'assortiment spécifique des allèles est également identique. Un compte-rendu détaillé des stades de la mitose est présenté dans Les grandes découvertes de la génétique 4-2.

MESSAGE

Au cours de la mitose, les molécules filles d'ADN (visibles sous la forme de chromatides sœurs) sont tirées vers les pôles opposés de la cellule, où elles forment deux noyaux identiques dans les cellules filles résultant de la division cellulaire.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 4-2

La mitose

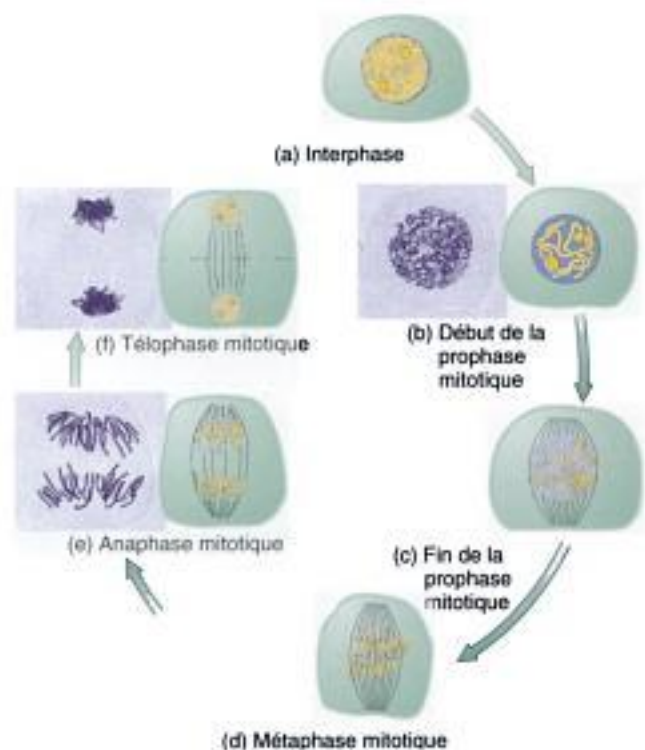
La mitose (M) est généralement la partie la plus brève du cycle cellulaire et représente approximativement 5 à 10 pour cent de la durée du cycle. La synthèse d'ADN a lieu durant la période S. G1 et G2 sont des phases intermédiaires entre S et M. L'ensemble G1, S et G2 constitue l'**interphase**, le temps qui sépare deux mitoses. (L'interphase était autrefois appelée « période de repos ». Pourtant, les cellules fourmillent d'activités durant cette période, la moindre n'étant pas la réplication de l'ADN). Les chromosomes ne sont pas visibles au cours de l'interphase (partie a de la figure) principalement en raison du fait qu'il sont dans un état relâché et sont entremêlés les uns aux autres dans un écheveau inextricable.

Pour en faciliter l'étude, les biologistes divisent la mitose en quatre stades appelés **prophase**, **métaphase**, **anaphase** et

télophase. Il faut pourtant insister sur le fait que toute division nucléaire est un processus dynamique auquel nous imposons des états arbitraires seulement pour notre commodité.

La prophase. Le début de la mitose correspond au moment où les chromosomes deviennent distincts pour la première fois (partie b). Ils se raccourcissent progressivement au cours d'un processus de contraction ou de condensation et s'organisent en une série de spirales et de tours. L'enroulement produit des structures qui sont plus faciles à déplacer. Lorsque les chromosomes deviennent visibles, ils apparaissent comme des doubles-brins, chaque chromosome étant constitué de deux moitiés longitudinales, les chromatides sœurs (partie c). Les deux chromatides formés par un chromosome contiennent chacune une des molécules répliquées d'ADN. La réplication étant semi-conservative, ces molécules d'ADN répliquées sont chacune « à moitié neuves et à moitié anciennes », c'est-à-dire que dans chaque double hélice, l'un des brins nucléotidiques vient d'être polymérisé. Néanmoins, malgré le fait qu'elles soient à moitié neuves et à moitié anciennes, les molécules d'ADN répliquées sont chimiquement identiques. Ces chromatides sœurs sont liés ensemble au niveau d'une région appelée le **centromère**. À ce stade, le centromère a déjà subi une division en une paire de centromères frères. Les nucléoles — les grosses structures sphériques intranucléaires qui contiennent les composants ribosomiques — disparaissent durant la prophase. La membrane nucléaire commence à se rompre, permettant au nucléoplasme et au cytoplasme de former un même espace.

La métaphase. A ce stade, le **fuseau nucléaire** devient proéminent. Le fuseau est une sorte de cage à oiseau qui se forme dans la zone nucléaire. Il est constitué d'une série de fibres protéiniques parallèles dirigées vers les deux pôles de la cellule. Ces fibres du fuseau sont des polymères d'une protéine appelée



NdT :

Des découvertes récentes montrent que ce sont en fait les kinétochores qui sont responsables du déplacement des chromosomes le long des fibres du fuseau.

Copyrighted material

tubuline. Les chromosomes migrent vers le plan équatorial de la cellule, où un centromère frère vient s'attacher à une fibre du fuseau émise à partir d'un pôle ; l'autre centromère frère se fixe à une autre fibre émise à partir de l'autre pôle (partie d).

L'anaphase. Ce stade débute lorsque les paires de chromatides sœurs se séparent, chacune migrant vers un pôle (partie e). Les centromères, qui apparaissent à présent clairement divisés, se séparent les premiers. Lorsque chaque chromatide se déplace, les deux bras semblent suivre leur centromère. Il en résulte un groupe de structures en V, avec les pointes des V dirigées vers les pôles.

La télophase. À présent, une membrane nucléaire se reforme autour de chaque noyau fils, les chromosomes se décondensent et les nucléoles réapparaissent — tout ceci reformant le noyau interphasique (partie f). Vers la fin de la télophase, le fuseau s'est dispersé et le cytoplasme a été divisé en deux par une nouvelle membrane cellulaire.

Dans chacune des cellules filles résultantes, la garniture chromosomique est identique à celle de la cellule originelle. Ce dont nous avons parlé sous le terme de *chromatides* reprend bien sûr son rôle de chromosomes à part entière.

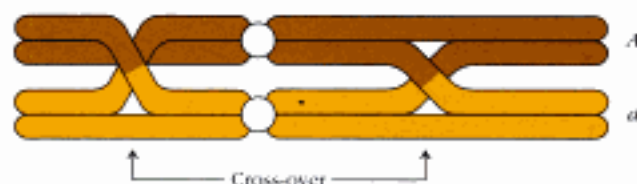
La division cellulaire sexuée

Bien que certaines bactéries aient une forme simple de reproduction sexuée (voir chapitre 9), il n'y a pas de division cellulaire spécialisée associée. Au contraire, la plupart des organismes peuvent se reproduire sexuellement et au cours de la partie sexuée du cycle vital, des divisions cellulaires spécialisées ont lieu. Chez les végétaux et les animaux, les divisions cellulaires sexuées produisent des ovules et des spermatozoïdes (gamètes) et chez les champignons, des spores sexuelles telles que les ascospores. Le point de départ de la division cellulaire sexuée est toujours une cellule diploïde appelée un **méiocyte** (figure 4-6). Chez la plupart des organismes complexes, tels que les animaux et les plantes à fleurs, les cellules de l'organisme sont normalement diploïdes et les méiocytes sont simplement une sous-population de cellules qui sont mises de côté pour la division sexuée — par exemple, celles que l'on trouve dans les ovaires et les testicules des animaux. Chez les organismes haploïdes, un méiocyte diploïde transitoire est formé au cours du cycle de reproduction normal (figure 4-7). Chez tous les Eucaryotes, les méiocytes se divisent deux fois, donnant lieu à quatre cellules haploïdes que l'on appelle une **tétrade**. Les deux divisions nucléaires qui accompagnent les deux divisions cellulaires sexuées sont appelées **méiose**.

Au contraire de la mitose dont le but est une propagation *conservative* d'un génotype, la méiose est un processus produisant de la *diversité*. Elle brasse les combinaisons alléliques, de sorte que si le méiocyte contient des paires d'allèles hétérozygotes (comme c'est le cas pour la plupart), les cellules qui représentent les **produits de la méiose** (c'est-à-dire les cellules des tétrades) contiendront de nombreuses combinaisons différentes de ces allèles. Les principaux moments où se déroule la méiose sont représentés dans les schémas des cycles vitaux de la figure 4-7.

La méiose est précédée par la phase de synthèse d'ADN dans le méiocyte diploïde. Ceci conduit au même résultat que la phase S avant une mitose diploïde — chaque chromosome dans les deux jeux est répliqué en une paire de chromatides sœurs, comme le montre précisément la figure 4-8. Juste avant la première division nucléaire, les chromosomes homologues s'apparient sur toute leur longueur, de sorte qu'à ce moment, pour chaque type de chromosome, il y a deux paires de chromatides sœurs juxtaposées, formant un faisceau de quatre chromatides également appelé une **tétrade**. La figure 4-9 présente une tétrade méiotique de chromosomes homologues. (Comparez cet empilement d'homologues avec la métaphase de la mitose, où au contraire, les chromosomes s'alignent côte à côte). Au stade de la tétrade, un processus remarquable se produit : des paires de chromatides *non sœurs* échangent des segments homologues d'ADN grâce à des cassures et des réunions de leurs bras au niveau de sites appelés **crossing-over**.

Lors de la première division de méiose, les centro-



mères agissent comme s'ils n'étaient pas encore divisés, bien que l'on sache que la répllication de l'ADN centromérique a eu lieu. Les fibres du fuseau émises par chaque pôle se fixent aux centromères homologues, qui peuvent alors gagner les pôles opposés avec les chromatides sœurs, (figure 4-9). Dans les deux cellules formées par la première division, le nombre de centromères est donc divisé par deux. Lors de la seconde division de méiose, les centromères se divisent et les fibres du fuseau se fixent chacune à une chromatide sœur, répartissant les deux membres de chaque paire entre les deux pôles. (figure 4-9). L'ensemble des deux divisions de méiose produit donc quatre cellules, chacune contenant le nombre haploïde de chromosomes. C'est inévitable car il y a un seul doublement du matériel génétique (phase S préméiotique) et deux divisions par deux du matériel génétique, qui se produisent au cours des deux divisions cellulaires. Comme nous l'avons vu, ce groupe de quatre cellules haploïdes (qui constituent les produits d'une seule méiose) est également appelé une **tétrade** ; le faisceau de quatre chromatides qui constitue une tétrade est réparti lors de la méiose entre ces quatre cellules. Dans les produits de la méiose, les chromatides sont à nouveau appelées **chromosomes**. La figure 4-9 montre que si un méiocyte diploïde est hétérozygote (par exemple A/a), alors la moitié des produits haploïdes de la méiose portera l'allèle A et l'autre moitié, l'allèle a. On dit que ces allèles **ségrégent** lors de la méiose car ils sont répartis entre différentes cellules haploïdes.

Voici certaines des caractéristiques essentielles de la méiose :

1. Il y a un cycle de répllication d'ADN mais deux cycles de division nucléaire (généralement deux

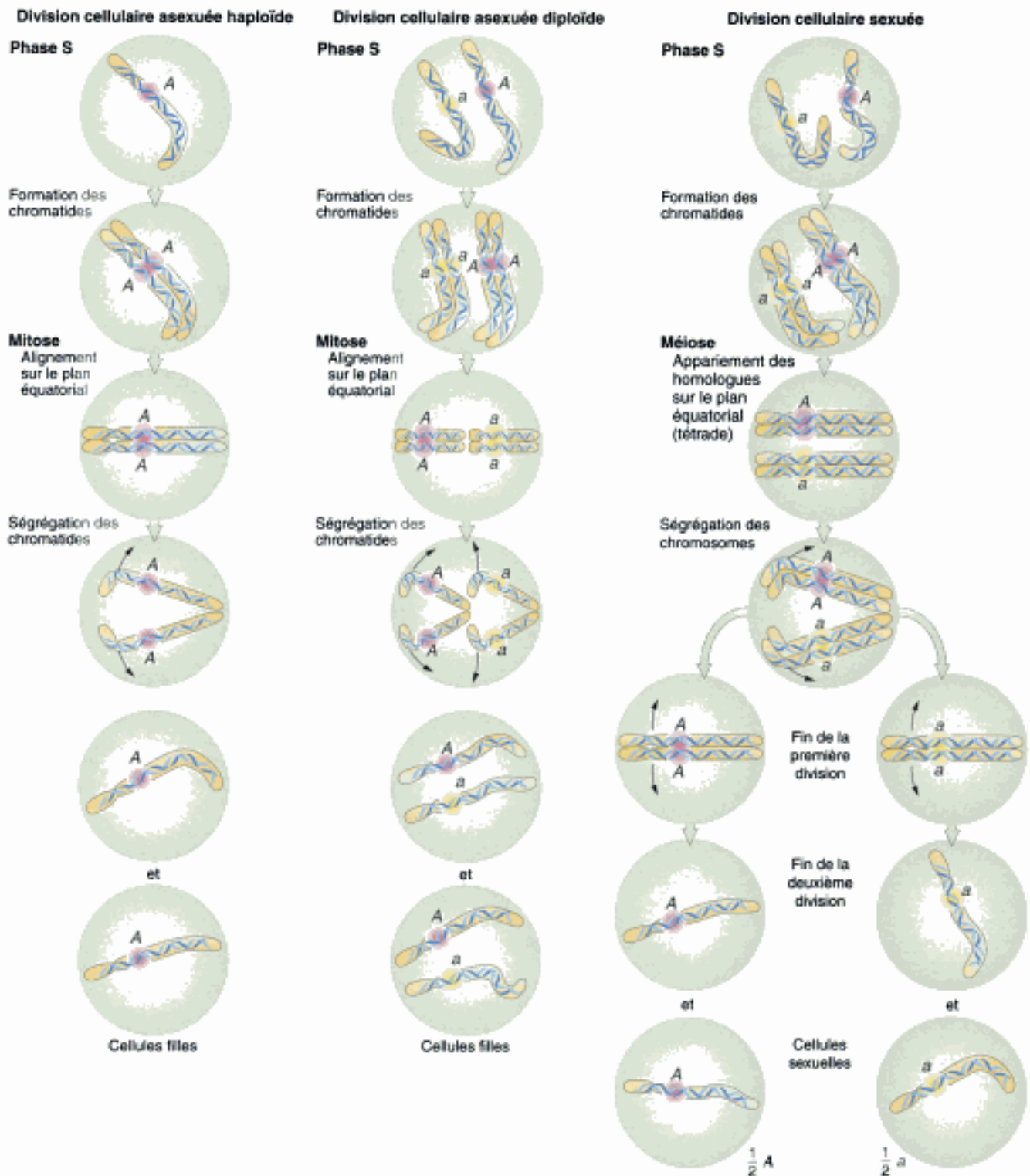


Figure 4-9 La transmission des gènes et de l'ADN au cours des divisions cellulaires sexuée et asexuée chez les Eucaryotes. La phase S et les principaux stades de la mitose et de la méiose sont présentés ici. Les schémas mettent l'accent sur le contenu en ADN de chaque cellule et de chaque chromosome. Les deux premiers dessins de chaque colonne présentent la réplication de l'ADN, qui se produit au cours de la phase S ; les autres dessins montrent la mitose ou la méiose. Les allèles *A* et *a* d'un gène sont utilisés pour montrer de quelle façon les génotypes sont transmis au cours de la division cellulaire. (Les symboles *A* et *a* seront utilisés pour représenter des allèles génériques dans le reste de ce chapitre).

cycles de division cellulaires également). Donc à la fin de la méiose, le nombre de chromosomes par cellule a diminué de moitié.

2. Lors de la première division méiotique, les centromères homologues sont tirés vers les pôles opposés de la cellule par les fibres du fuseau. Les fibres du fuseau provenant d'un pôle s'attachent au hasard à l'un ou l'autre des chromosomes d'une paire d'homologues (par leur centromères). De ce fait, les centromères hétérologues et les chromosomes dont ils font partie se combinent aléatoirement. Cette répartition indépendante produit de nombreux génotypes différents et est l'une des sources de la diversité produite par la méiose.
3. Au tout début de la première division méiotique, il y a un échange de matériel chromosomique entre les chromatides homologues, à la suite de crossing-over. Les crossing-over ont deux rôles. Premièrement, ils sont la deuxième source de diversité

génomique car deux chromatides ayant subi un crossing-over ne sont jamais exactement identiques. Deuxièmement, les événements d'échange qui se produisent entre des chromatides homologues dans la tétrade servent à maintenir cette tétrade jusqu'à ce que les homologues soient séparés, vers la fin de la première division. Ceci garantit la ségrégation correcte des chromosomes et empêche la formation de produits de méiose comportant un nombre anormal de chromosomes.

MESSAGE

Le processus de méiose s'engage dans des méiocytes diploïdes situés dans le tissu germinale et produit une série de cellules haploïdes de génotypes variés.

Les différences entre méiose et mitose sont résumées dans la figure 4-10. La méiose elle-même est décrite plus en détail dans Les grandes découvertes de la génétique 4-3.

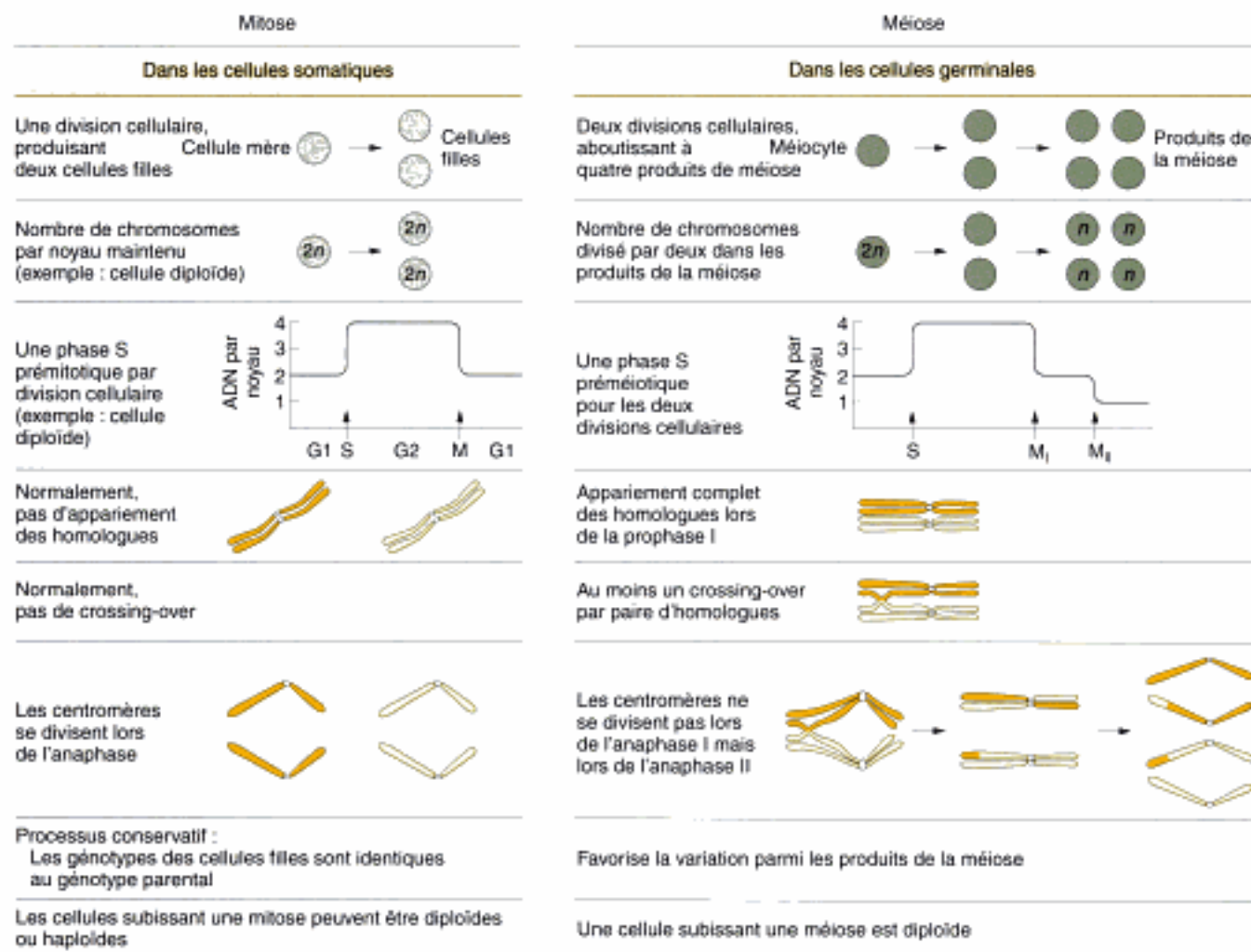


Figure 4-10 Comparaison des caractéristiques principales de la mitose et de la méiose.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 4-3

La méiose

La méiose comporte deux divisions nucléaires que l'on appelle **méiose I** et **méiose II**. Les événements de la méiose I sont assez différents de ceux de la méiose II et tous deux différent des étapes de la mitose. Chaque division méiotique est divisée formellement en prophase, métaphase, anaphase et télophase. Le plus complexe et le plus long de ces stades est la prophase I qui se divise à son tour en cinq étapes. Une fois encore, essayez d'envisager ces étapes s'enchaînant de façon dynamique sans frontières nettes.

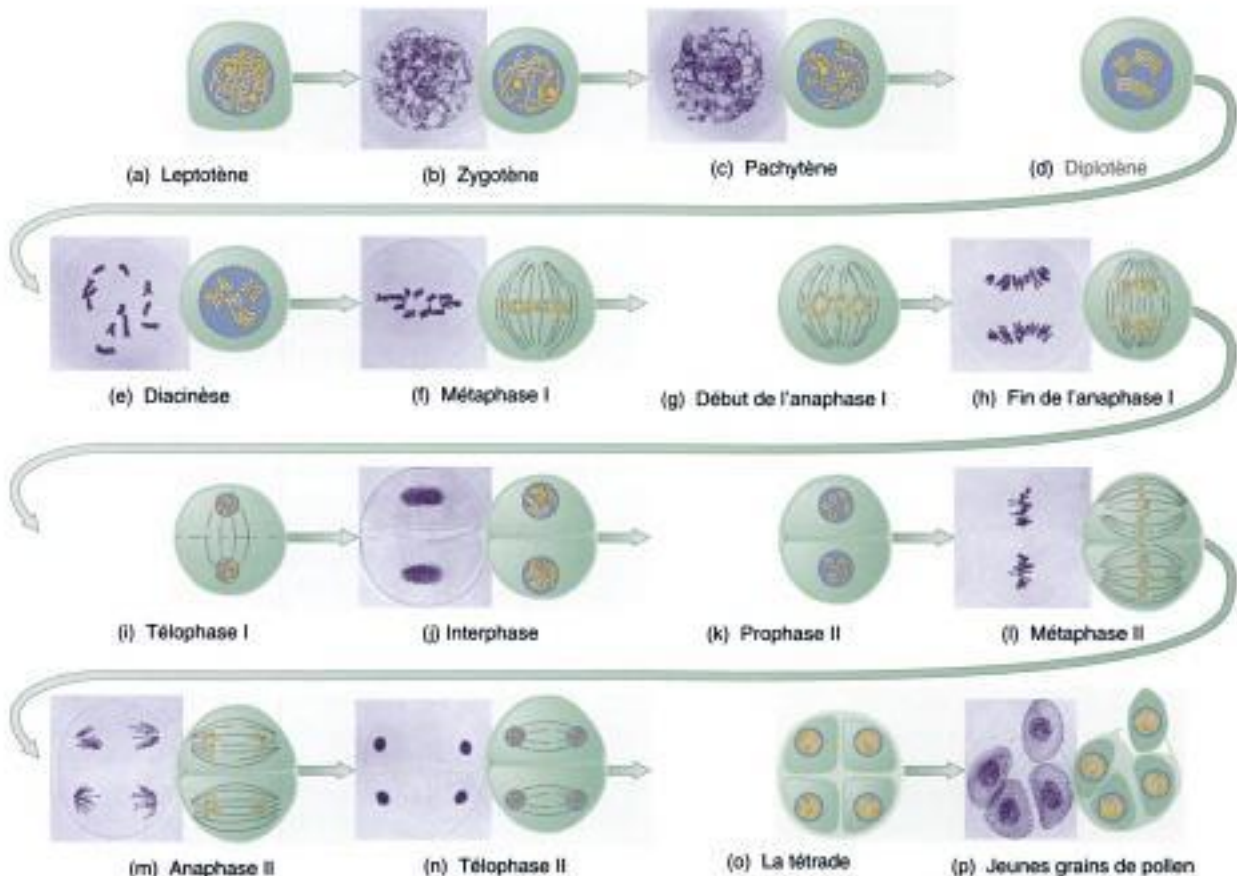
La prophase I. Pour des raisons de commodité, la prophase I est divisée en plusieurs stades. Une grande partie de ce qui se produit au cours des différents stades de la prophase I est un processus continu mettant en jeu la condensation des chromosomes et la stabilisation des associations de chromosomes homologues au sein d'une tétrade, de sorte que la tétrade peut subir une ségrégation égale lors de l'anaphase.

Leptotène. Les chromosomes deviennent visibles à ce stade sous la forme de longs fils fins individuels (voir partie a de la figure). Bien qu'ils semblent uniques, ils sont en réalité composés de paires de molécules d'ADN répliquées. Le processus de

contraction des chromosomes se poursuit au cours du stade leptotène et pendant la totalité de la prophase. Au cours du stade leptotène, de petits renflements appelés **chromomères** apparaissent le long de chacun des chromosomes, ce qui donne à ceux-ci une apparence de collier de perles.

Zygotène. C'est une période d'appariement actif des chromosomes (partie b) qui permet de voir que la garniture chromosomique du mélocyte est en fait constituée de deux jeux complets de chromosomes. Chaque chromosome a donc un partenaire avec lequel s'apparier et tous deux s'apparient progressivement sur toute leur longueur. Il faut noter que cet appariement n'a pas lieu au cours de la mitose, car pendant celle-ci tous les chromosomes sont répartis au hasard dans le plan équatorial, sans se soucier de leurs homologues.

Comment des homologues (les membres d'une paire de chromosomes) peuvent-ils se retrouver au cours du stade zygotène ? La réponse probable est que les extrémités des chromosomes, les **télomères**, sont ancrés dans la membrane nucléaire. Il est probable que les télomères des homologues soient proches, de sorte que les deux homologues peuvent s'apparier précisément sur toute leur longueur. Bien que le mécanisme de cet appariement ne soit pas parfaitement compris, on sait qu'il implique une structure composée de protéines et d'ADN appelée **complexe synaptonémal**, que l'on trouve toujours en sandwich entre des homologues au cours de leur appariement. La microphotographie de gauche présente plusieurs vues de





(a)

Complexes synaptonémaux. (a) Chez *Hyalophora cecropia*, un papillon de nuit, le nombre normal de chromosomes pour un mâle est de 62, donnant lieu à 31 complexes synaptonémaux. Chez l'individu présenté ici, un chromosome (au centre) est représenté trois fois. Un tel chromosome est qualifié de trivalent. L'ADN est arrangé en boucles régulières autour du complexe synaptonéma. La structure noire et dense est le nucléole. (b) Complexe synaptonéma classique chez *Lilium tyrinum*. Remarquez (à droite) les deux éléments latéraux du complexe synaptonéma et également (à gauche) un chromosome non apparié, présentant une partie centrale correspondant à l'un des éléments latéraux. [Parties (a) et (b) aimablement communiquées par Peter Moens].



(b)

complexes synaptonémaux. La partie a montre les 62 chromosomes d'un bombyx appariés en 31 complexes. Dans la partie b se trouve à gauche un chromosome non apparié (remarquez la ligne noire qui passe en son centre). Cette ligne noire est répétée deux fois dans le complexe synaptonéma présenté à droite, représentant les centres des deux homologues alignés l'un avec l'autre.

Pachytène. Ce stade est caractérisé par des chromosomes épais, complètement appariés (partie c). Le nombre de paires homologues de chromosomes dans le noyau est donc égal au nombre n . Les nucléoles sont souvent bien visibles au cours du stade pachytène. Les chromomères, semblables à des perles, s'alignent précisément sur les homologues appariés, produisant un aspect différent pour chaque paire.

Diploène. Bien que chaque homologue apparaisse sous la forme d'un seul fil au cours du stade leptotène, l'ADN a en fait déjà été répliqué au cours de la phase S prémeiotique. Ceci devient visible au cours du stade diploène, sous la forme d'une duplication longitudinale de chaque homologue apparié (partie d). Comme dans la mitose, les sous-unités divisées sont appe-

lées *chromatides sœurs*. Puisque chaque membre d'une paire d'homologues produit deux chromatides sœurs, la structure appariée consiste alors en quatre chromatides homologues, la tétrade. Au stade diploène, l'appariement entre les homologues devient moins fort : ces derniers semblent en fait se repousser les uns les autres et à mesure qu'ils se séparent, des structures en forme de croix apparaissent entre les chromatides non sœurs. C'est ce que l'on appelle des **chiasm**. Chaque groupe d'homologues comporte généralement un ou plusieurs chiasm. Les chiasm sont les manifestations visibles d'événements appelés **crossing-over** qui se sont produits plus tôt, probablement au cours des stades zygotène ou pachytène. Les crossing-over représentent l'une des différences principales entre la méiose et la mitose — les crossing-over sont rares durant la mitose. Un crossing-over comporte une cassure précise, un échange et une réunion de fragments entre deux chromatides non sœurs. Les études réalisées sur des lignées anormales d'organismes qui subissent des crossing-over de manière très inefficace ou qui n'en subissent pas du tout, montrent une perturbation sévère des événements ordonnés qui répartissent les

chromosomes entre les cellules filles lors de la méiose. La formation de crossing-over aide donc à l'évidence à déterminer de quelle façon se comportent des homologues appariés, et la présence d'au moins un crossing-over par paire est généralement essentielle pour une ségrégation correcte. Les crossing-over jouent un autre rôle intéressant. Comme nous le verrons au chapitre 5, ils permettent de former de nouvelles combinaisons alléliques, ce qui constitue une source importante de variation génétique dans les populations.

Diacynèse. Ce stade (partie e) ne diffère que très légèrement du stade diplotène à l'exception d'une contraction supplémentaire des chromosomes. Vers la fin de la diacynèse, les longs filaments chromosomiques de l'interphase ont été remplacés par des unités compactes qui sont bien plus manœuvrables.

La métaphase I. La membrane nucléaire et les nucléoles ont disparu lors de la métaphase I et chaque paire d'homologues prend position sur le plan équatorial (partie f). À ce stade de la méiose, bien qu'il soit probable que la réplication de l'ADN ait eu lieu au niveau des centromères, ceux-ci n'apparaissent pas divisés. Les deux centromères frères agissent donc comme un seul. Cette absence de division apparente représente une différence essentielle avec la mitose. Les deux centromères non frères se fixent aux fibres du fuseau émises à partir des pôles opposés.

L'anaphase I. Comme pour la mitose, l'anaphase commence lorsque les chromosomes se dirigent vers les pôles. Les paires de chromatides sœurs se dirigent vers les pôles opposés (parties g et h).

La télophase I. Cette télophase (partie i) et l'« interphase » qui la suit, appelée **intercinèse** (partie j) ne sont pas universelles. Chez de nombreux organismes, ces stades n'existent pas, aucune membrane nucléaire ne se reforme et les cellules s'engagent directement dans la méiose II. Dans d'autres organismes, la télophase I et l'intercinèse sont de courte durée. Les chromosomes s'allongent, deviennent diffus et la membrane nucléaire se reforme. Dans tous les cas, il n'y a jamais de synthèse d'ADN à ce moment et l'état génétique des chromosomes ne change pas.

La prophase II. La présence du nombre haploïde de paires de chromatides sœurs dans l'état condensé caractérise la prophase II (partie k).

La métaphase II. Les paires de chromatides sœurs s'arrangent sur le plan équatorial au cours de la métaphase II (partie l). À ce moment-là, les chromatides se dissocient souvent partiellement l'un de l'autre au lieu de rester accolés l'un à l'autre comme au cours de la mitose.

L'anaphase II. Les centromères se divisent et les chromatides sœurs gagnent les pôles opposés à l'aide des fibres du fuseau (partie m).

La télophase II. Les noyaux se reforment autour des chromosomes situés au niveau des pôles (partie n).

Les quatre produits de la méiose sont présentés dans la partie o. Dans les anthères d'une fleur, les quatre produits de la méiose se développent en grains de pollen ; ils sont présentés dans la partie p. Chez les autres organismes, la différenciation produit d'autres sortes de structures à partir des produits de la méiose, telles que les spermatozoïdes chez les animaux.

LES MODES DE TRANSMISSION DES GÈNES INDIVIDUELS

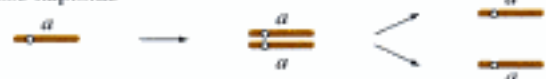
Les modes de transmission des gènes individuels sont dictés par les déplacements hautement programmés des chromosomes lors de la mitose et de la méiose.

Revenons tout d'abord à la mitose. Nous utiliserons les allèles A et a comme les allèles « type » d'un même gène. Les cellules haploïdes peuvent être de génotype A ou a et les diploïdes peuvent être homozygotes, A/A ou a/a , ou hétérozygotes, A/a . Chacun des chromosomes étant répliqué fidèlement, les génotypes des cellules filles *doivent* être identiques à celui de la cellule mère. Nous pouvons représenter ces divisions de la façon suivante :

Mitose haploïde



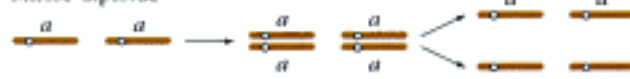
Mitose haploïde



Mitose diploïde



Mitose diploïde

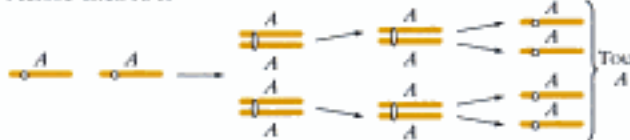


Mitose diploïde



Revenons maintenant au résultat de la méiose pour les différents génotypes. La logique voudrait que la méiose chez des méiocytes *homozygotes* ne puisse produire qu'un seul génotype dans les quatre produits haploïdes de la méiose (la tétrade) comme on le voit ci-dessous :

Méiose chez A/A



Méiose chez a/a

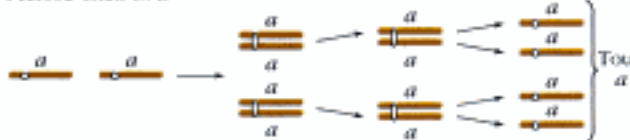
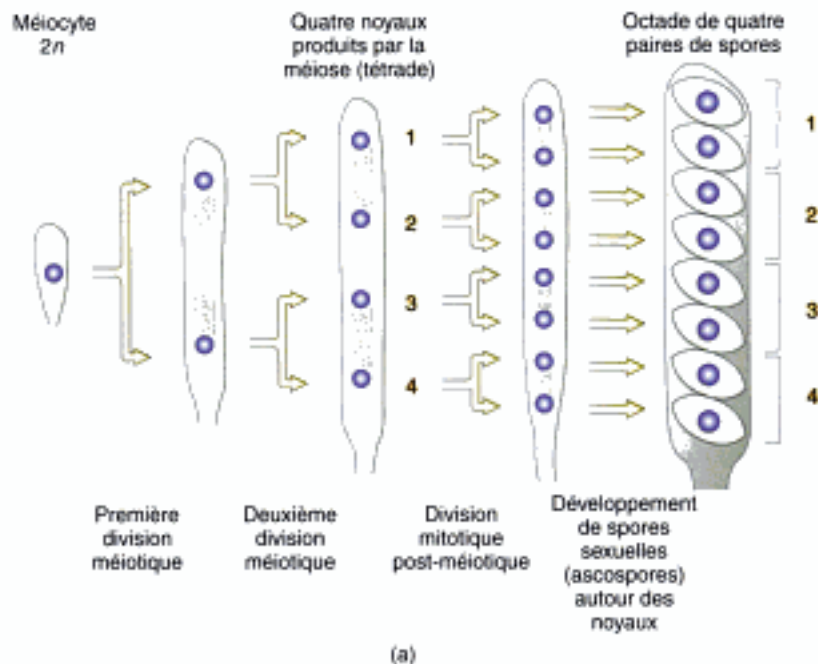
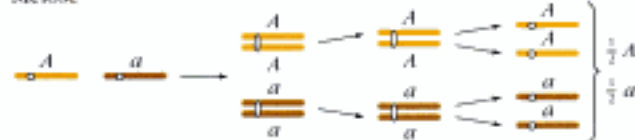


Figure 4-11 Certains champignons constituent d'excellents systèmes d'étude de la ségrégation des allèles lors de la méiose. (a) Chez le champignon *Neurospora* (moisissure rose du pain), la méiose engendre quatre noyaux (une tétrade) qui subissent alors une mitose pour produire une octade. Puisque les divisions nucléaires méiotiques et mitotiques se produisent toutes dans une longue cellule mince, les noyaux ne glissent jamais les uns par-dessus les autres ; les produits sont donc alignés. (b) Un méiocyte *A/a* subit une méiose (et après une mitose), ce qui aboutit à un nombre égal de produits *A* et *a*, et fournit une démonstration directe du principe de ségrégation égale des allèles lors de la méiose (première loi de Mendel). (Remarque : la méiose illustrée n'a aucun crossing-over dans la région comprise entre le centromère et le gène étudié. Un crossing-over à cet endroit changerait la disposition des ascospores dans l'asque mais ne modifierait pas le rapport 1 : 1 des allèles dans les produits de la méiose).



Toutefois, à partir d'un méiocyte de génotype *A/a*, la méiose produit quatre cellules haploïdes, la moitié étant *A* et la moitié *a* comme nous l'avons vu dans la figure 4-9.

Méiose

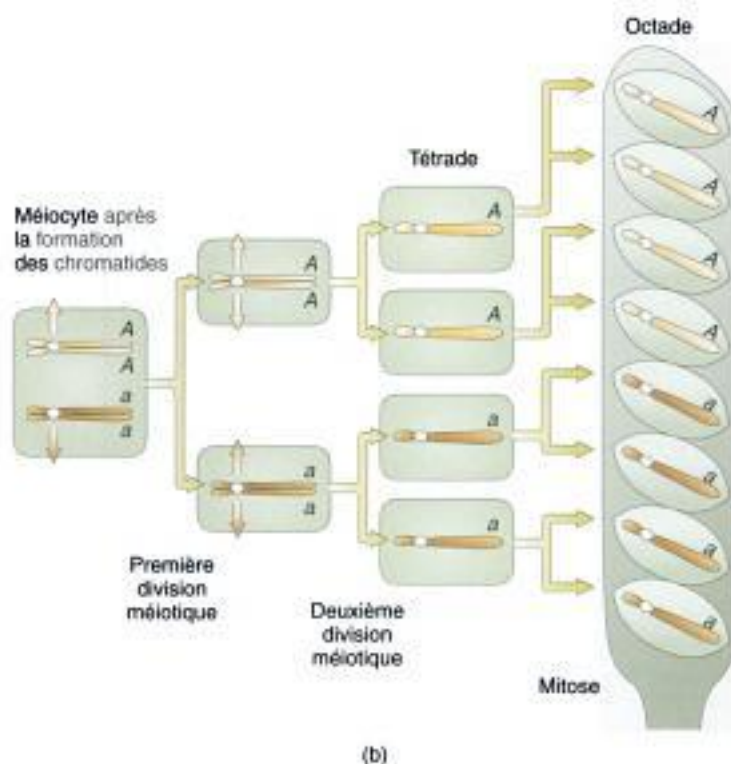


La raison de cette répartition est que dans le méiocyte *A/a*, le chromosome portant *A* produit une paire de chromatides sœurs *A/A* et que le chromosome homologue produit une paire *a/a*. Ces quatre « copies » du gène se retrouvent dans les quatre cellules produites par la méiose. Ce résultat, que l'on appelle **ségrégation égale**, a été observé pour la première fois par Gregor Mendel au dix-neuvième siècle (voir Les grandes découvertes de la génétique 4-4). On l'appelle parfois première loi de Mendel. La ségrégation 1 : 1 n'est pas affectée par des crossing-over.

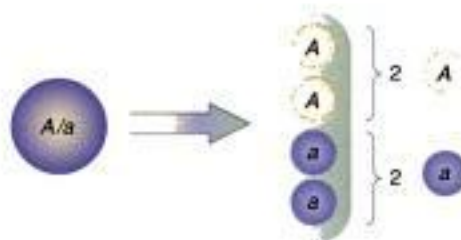
MESSAGE

La méiose chez un méiocyte hétérozygote engendre des produits haploïdes (par exemple des gamètes) dont une moitié porte un des allèles et l'autre moitié l'autre allèle ; par exemple,
 $A/a \rightarrow 1/2 A \text{ et } 1/2 a$.

Chez les animaux et les végétaux, il est impossible d'isoler une tétrade méiotique (les quatre cellules qui sont les produits de la méiose chez un *seul* méiocyte). Les ovules et les spermatozoïdes produits sont le résultat de méioses dans *de nombreux* méiocytes. Mais la loi de la ségrégation égale est toujours valable, donc sur 1000 spermatozoïdes et ovules produits par de nombreux méiocytes chez un hétérozygote *A/a*, 500 seront de génotype *A* et 500 seront *a*. (L'échantillonnage peut bien sûr présenter des variations par rapport à ce rapport 1 : 1). Toutefois, certains champignons et certaines algues ont des cycles sexuels qui permettent d'isoler les quatre produits d'une seule méiose. Deux champignons parentaux (disons de génotypes *A* et *a*) s'unissent sexuellement pour former un méiocyte *transitoire* qui sera *A/a* (consulter la figure 4-7 pour le cycle sexuel des champignons). La méiose se déroule et les produits qu'elle engendre sont des **spores sexuelles**. Les quatre spores sexuelles qui représentent les produits d'une seule méiose sont maintenues ensemble sous la forme d'une tétrade. Chez la levure (et d'autres champignons ascomycètes) la tétrade est contenue dans un sac appelé **asque**. Les asques de levure issues d'une paire parentale de ce type contiendront toujours 2 spores sexuelles *A* et 2 spores *a*. Chez l'ascomycète filamenteux *Neurospora crassa* et les champignons apparentés, les quatre noyaux haploïdes résultant de la méiose subissent une division mitotique produisant 8 spores sexuelles (une **octade**). Dans ce cas, la loi de la ségrégation égale serait exprimée dans l'asque par 4*A* et 4*a* (voir figure 4-11). La figure 4-12 montre la différence



Ségrégation 1:1 dans les produits d'un seul méiocyte



Ségrégation 1:1 dans des produits de méiose pris au hasard

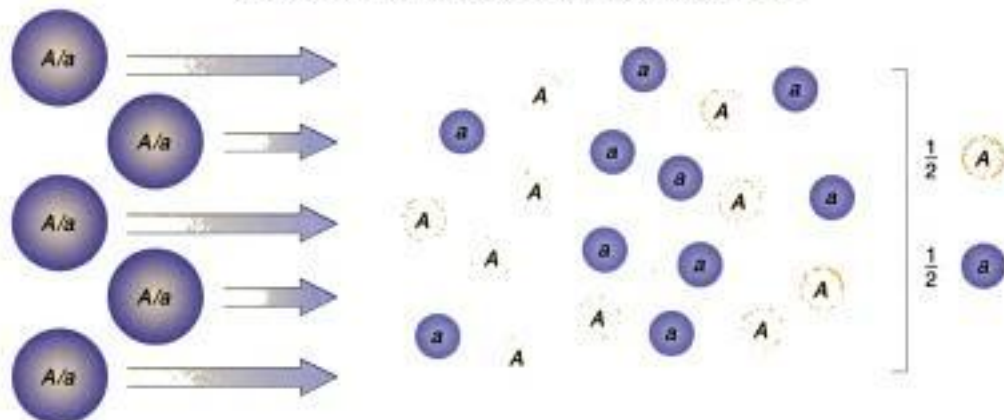


Figure 4-12 Représentation des différentes façons dont se manifeste la loi de ségrégation égale lors de l'analyse d'un seul méiocyte, comme dans les tétrades de champignons, ou de produits de méiose pris au hasard parmi de nombreux méiocytes identiques comme dans l'analyse des ovules et des spermatozoïdes chez les animaux.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 4-4

Mendel, l'inventeur du concept de gène

Dans son article de 1866 intitulé « Expériences sur l'hybridation de végétaux », le moine morave Gregor Mendel définit les principes de l'ensemble de la génétique en postulant pour la première fois l'existence d'entités héréditaires distinctes, que nous appelons maintenant des gènes. Le but de l'étude de Mendel était résumé par l'introduction de son article :

La fécondation artificielle entreprise sur des plantes d'ornement pour obtenir des variants de couleurs nouvelles est à l'origine des expériences décrites ici. L'étonnante régularité avec laquelle la même forme hybride réapparaît pour n'importe quelle fécondation réalisée entre les deux mêmes espèces a suggéré d'autres expériences pour suivre le développement des hybrides et de leur descendance.



Gregor Mendel (Musée Morave, Brno)

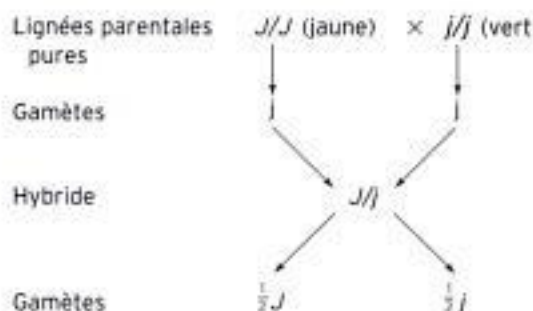
Les expériences de Mendel débutèrent avec des lignées pures de petits pois qui présentaient des phénotypes contrastés pour certains caractères. Par exemple, deux lignées produisaient des graines jaunes pour l'une et des graines vertes pour l'autre, présentant donc des phénotypes de couleur de graines clairement distincts. En croisant ces lignées, Mendel obtint des graines « hybrides » qui étaient jaunes. Toutefois, après culture et auto-pollinisation des hybrides, 3/4 des graines de la génération suivante étaient jaunes et 1/4 étaient vertes.

Sur les 3/4 des graines jaunes, 2/4 étaient exactement comme les hybrides jaunes de départ tandis que 1/4 étaient de lignée pure, comme la lignée jaune d'origine.

Pour expliquer ces rapports et d'autres rapports semblables obtenus à partir de paires de phénotypes distincts pour d'autres caractères, Mendel formula les hypothèses suivantes :

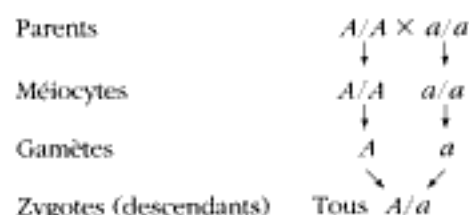
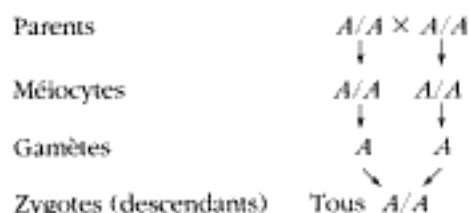
1. La différence entre le jaune et le vert résulte de différences entre des déterminants héréditaires discrets qu'il appela *facteurs* (ce que nous appelons aujourd'hui des « gènes »).
2. Les facteurs existent par paires, une paire pour chaque caractère.
3. Lors de la formation des gamètes, les membres d'une paire se séparent, chacun étant distribué dans un gamète sur deux.
4. Les gamètes mâles et femelles fusionnent au hasard.
5. La couleur jaune est dominante sur la couleur verte, qui est récessive.

Dans son analyse, il représenta les gènes par des lettres et ce fut le début d'une pratique encore en usage aujourd'hui. En nommant *J* le gène déterminant la couleur jaune et *j* le gène déterminant la couleur verte, l'analyse des croisements était présentée ainsi :



La génération suivante, après une fécondation au hasard des femelles hybrides par des gamètes mâles hybrides, aurait conduit au résultat suivant :

Hidden page



Les croisements impliquant des hétérozygotes sont légèrement plus complexes, mais la loi de ségrégation égale peut être utilisée pour prédire les proportions des gamètes. Les tableaux suivants montrent de quelle façon les descendants sont produits dans les trois croisements différents mettant en jeu des hétérozygotes. Les rapports des gamètes produits par les deux parents sont indiqués en haut et sur le côté des tableaux et le tableau lui-même décrit les cellules résultant de la fusion des gamètes (les zygotes).

$A/a \times a/a$

		Méiocyte A/a	
		\swarrow	\searrow
Gamètes		$\frac{1}{2} A$	$\frac{1}{2} a$
Méiocyte $a/a \rightarrow$	Tous a	$\frac{1}{2} A/a$	$\frac{1}{2} a/a$

Rapport des phénotypes chez la descendance : **1/2 dominants, 1/2 récessifs**

$A/a \times A/A$

		Méiocyte A/a	
		\swarrow	\searrow
Gamètes		$\frac{1}{2} A$	$\frac{1}{2} a$
Méiocyte $A/A \rightarrow$	Tous A	$\frac{1}{2} A/A$	$\frac{1}{2} A/a$

Rapport des phénotypes chez la descendance : **tous dominants** (l'écriture du génotype $A/-$ signifie que le deuxième allèle peut être soit A , soit a).

$A/a \times A/a$

		Méiocyte A/a	
		\swarrow	\searrow
Gamètes		$\frac{1}{2} A$	$\frac{1}{2} a$
Méiocyte $A/a \rightarrow$	$\frac{1}{2} A$	$\frac{1}{4} A/A$	$\frac{1}{4} A/a$
	$\frac{1}{2} a$	$\frac{1}{4} A/a$	$\frac{1}{4} a/a$

Rapport des phénotypes chez la descendance :

$\frac{3}{4}$ **dominants ($A/-$)**

$\frac{1}{4}$ **récessifs (a/a)**

Remarquez que sous le rapport *phénotypique* 3 : 1 de ce croisement se cache un rapport *génétypique* de 1 : 2 : 1 correspondant à A/A : A/a : a/a .

Rapport génétique Rapport phénotypique

$\frac{1}{4} A/A$	$\left. \begin{array}{l} \frac{3}{4} \text{ dominants } A/- \\ \frac{1}{4} \text{ récessifs } a/a \end{array} \right\}$
$\frac{1}{2} A/a$	
$\frac{1}{4} a/a$	

Pour analyser un seul gène, on peut réaliser un **croisement-test** (*testcross*) afin de déduire si un individu présentant le phénotype dominant est homozygote (A/A) ou hétérozygote (A/a). L'individu de phénotype dominant $A/-$ (c'est-à-dire soit A/A soit A/a) est croisé avec un homozygote récessif a/a (appelé **souche-test** dans ce contexte). S'il y a un rapport 1 : 1 entre les phénotypes dominant et récessif dans la descendance, on peut alors en déduire que le croisement a été du type $A/a \times a/a$ et que la souche étudiée devait être hétérozygote. Si toute la descendance présente le phénotype dominant, alors la souche devait être homozygote A/A et le croisement-test était $A/A \times a/a$, donnant des descendants de génotype A/a .

Par exemple chez les souris, la différence entre un pelage noir et un pelage marron est déterminée par une paire d'allèles, B (pour *black* : noir en anglais) et b (pour *brown* : marron en anglais), où B est dominant sur b . On réalise un croisement-test entre un mâle noir d'ascendance inconnue et une femelle marron. Le croisement aboutit à huit petits : trois sont noirs et cinq sont marron. Ce résultat nous indique que le mâle devait être homozygote B/b car les souriceaux marron ont dû recevoir un allèle b à la fois de leur mère et de leur père. Le croisement a donc été $B/b \times b/b$. La loi de Mendel sur la ségrégation égale nous laisserait évidemment attendre des nombres égaux de descendants noirs et marron (4 : 4) mais un écart tel que 3 : 5 peut survenir dans un échantillon d'une si petite taille.

Les animaux et les végétaux utilisés pour l'analyse génétique sont souvent maintenus à l'état de **lignées pures** (également appelées *souches* ou *lignées de croisement pures*). Il s'agit de populations qui, lorsqu'on les croise entre elles ou lorsqu'elles se reproduisent par autofécondation, présenteront toujours le même phénotype pour un caractère particulier. De telles lignées doivent être homozygotes. Les génotypes A/A et a/a sont des génotypes de lignées pures mais A/a n'en est pas un.

Utiliser des croisements pour déduire les génotypes, la dominance et la récessivité. En obtenant les rapports ci-dessus, nous avons utilisé des génotypes parentaux connus impliquant des allèles dont la dominance ou la récessivité était connue, pour réaliser des prédictions sur les rapports phénotypiques attendus dans la descendance. Pour déduire les génotypes et la dominance lorsqu'ils sont inconnus, il nous faut effectuer les croisements appropriés. Mendel a été le premier à mettre en place ce type de procédure expérimentale. Nous pouvons donc utiliser un exemple de son travail, concernant la couleur de graines de pois. La procédure de croisement est indiquée dans la figure 4-13. Dans cette procédure, le symbole P désigne par convention la génération **parentale** et F_1 et F_2 correspondent aux **première et deuxième générations filiales**.

Pour explorer la génétique sous-tendant le caractère de la couleur des graines, nous croisons deux lignées pures de pois, l'une jaune et l'autre verte, pour produire la F_1 . Une F_2 est alors obtenue par autofécondation ou par croisement entre des plantes de la F_1 . Nous observons que toutes les plantes de la F_1 présentent des graines jaunes et que parmi les plantes de la F_2 , 3/4 ont des graines jaunes et 1/4 des graines vertes, soit un rapport phénotypique 3 : 1. Deux déductions peuvent être faites à partir de ces résultats :

1. Le phénotype vert disparaît dans la F_1 mais réapparaît dans la F_2 . Nous pouvons donc en déduire que le jaune devait dominer le vert chez la F_1 . Jaune est donc le phénotype dominant et vert le phénotype récessif.
2. Puisque nous savons qu'un rapport 3 : 1 dans la F_2 révèle un croisement entre deux hétérozygotes pour un même gène, nous pouvons en déduire que les deux phénotypes jaune et vert sont déterminés par les allèles d'un même gène et que les plantes de la F_1 doivent avoir un phénotype J/j , où J est un allèle dominant déterminant la couleur jaune et j un allèle récessif déterminant la couleur verte. (Remarque : un croisement entre deux hétérozygotes pour un même gène est parfois appelé **croisement monohybride**).

MESSAGE

L'analyse génétique fonctionne en deux sens :

1. En croisant des parents de génotype et de relation de dominance connus pour obtenir des types de descendants spécifiques
2. En observant les phénotypes et les rapports des descendants pour en déduire les génotypes des parents et les relations de dominance.

Cette technique analytique a révélé le génotype et les relations de dominance pour les caractères de nombreux organismes. D'autres exemples sont indiqués plus loin.

Chez la levure (haploïde), les colonies sont normalement blanches. Cependant, une souche rare est apparue, qui possède des colonies roses. Un croisement a été effectué entre une souche blanche et une souche rose et

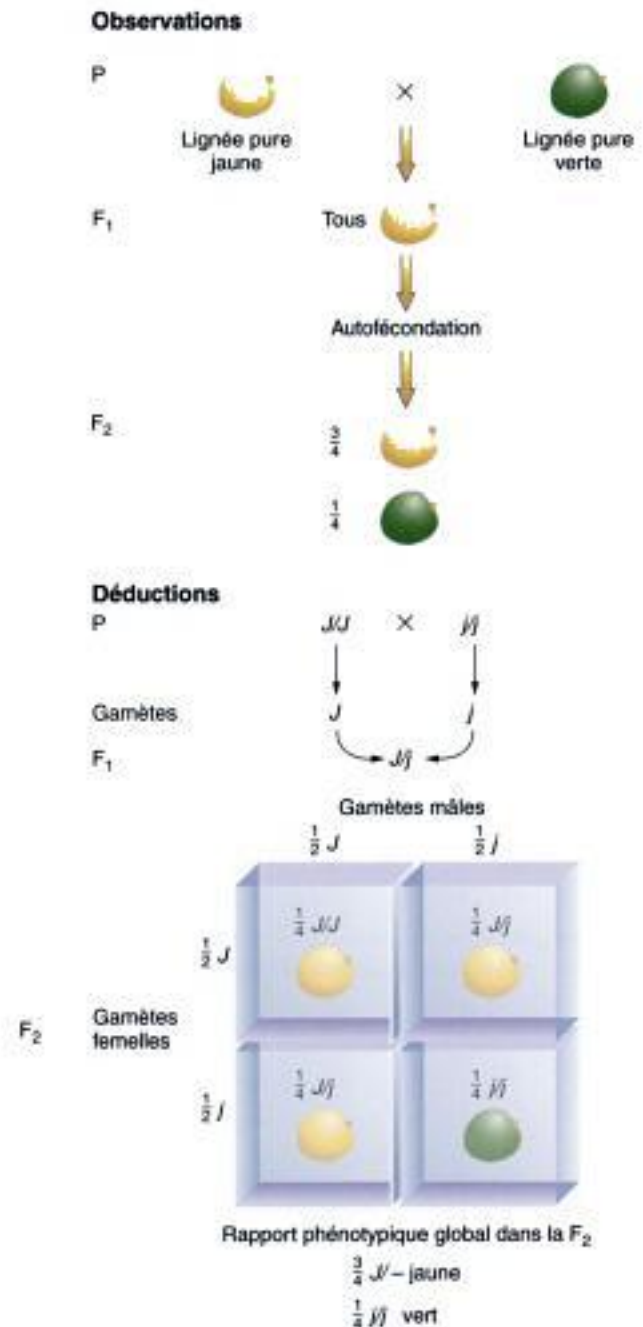
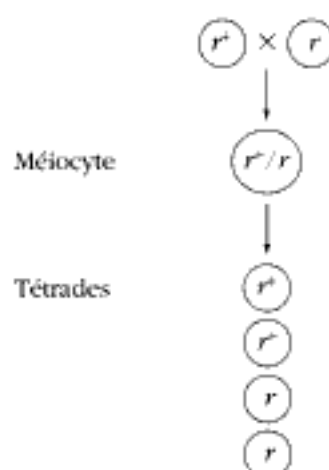


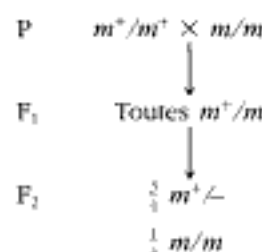
Figure 4-13 Utilisation de lignées pures pour déduire les génotypes ainsi que la dominance et la récessivité.

une analyse de tétrades a été réalisée. Toutes les tétrades présentaient 2 spores blanches et 2 spores roses. Ce résultat montre la ségrégation 1 : 1 attendue pour des méiocytes hétérozygotes. Nous pouvons donc en déduire que la différence de phénotype a été déterminée par deux allèles d'un gène. Posons r = rose et r^+ = blanc (type sauvage). Alors le croisement a dû être le suivant :



Les drosophiles ont normalement d'assez grandes ailes (type sauvage). En laboratoire, on découvrit une mouche avec des ailes miniatures. Lorsqu'on croisa une mouche de cette sorte avec une mouche de type sauvage, toutes les mouches de la F_1 présentaient de grandes ailes. Lorsqu'on laissa les mouches de la F_1 s'accoupler entre elles, produisant ainsi une F_2 , on dénombra 400 mouches, dont 305 avaient des ailes normales et 95 des ailes miniatures. Ce résultat nous indique que :

1. Les deux phénotypes doivent être déterminés par deux allèles d'un même gène car le rapport de 3 : 1 observé pour la F_2 est basé sur une ségrégation méiotique 1 : 1 chez un hétérozygote (c'est-à-dire un rapport correspondant à un *seul* gène).
2. Les grandes ailes doivent être dominantes sur les ailes miniatures car toutes les mouches de la F_1 avaient de grandes ailes.
Si l'on pose m = miniature et m^+ = grande, le croisement a dû être :



Les gènes autosomiques et les gènes liés au sexe

Le sexe est déterminé par les chromosomes chez de nombreux animaux, tels que les humains et les drosophiles et chez les plantes dioïques (celles dont les organes mâles et femelles se trouvent sur des plantes séparées). Dans tous ces cas, il est nécessaire de distinguer les **chromosomes sexuels** (ceux qui sont impliqués dans la détermination du sexe) et les **autosomes** (les autres chromosomes présents chez les deux sexes). Il y a généralement une seule paire de chromosomes sexuels. Chez les femmes, les chromosomes sexuels sont une paire de chromosomes X et chez les hommes il y a un chromosome X et un Y. Si l'on utilise la lettre A pour désigner un autosome, une femme sera représentée par la combinaison 44A + XX et un homme par 44A + XY. L'ensemble des chromosomes humains est présenté dans la figure 4-14. Chez les femmes, la paire de chromosomes X homologues subit une ségrégation lors de la méiose, de sorte que tous les ovules possèdent un seul X. Chez les hommes, le X et le Y sont homologues seulement pour une petite région située à une extrémité, mais lors de la méiose ils s'apparient au niveau de cette région et subissent une ségrégation égale de sorte que 1/2 des spermatozoïdes portent un X et 1/2 un Y. Lorsque ces spermatozoïdes fusionnent au hasard avec les ovules portant les X, les descendants sont à 50 pour cent des garçons (XY) et à 50 pour cent des filles (XX).

		Spermatozoïdes	
		Gamètes $\frac{1}{2}$ X	$\frac{1}{2}$ Y
Ovules \rightarrow	Tous X	$\frac{1}{2}$ X/X (fille)	$\frac{1}{2}$ X/Y (garçon)

L'égalité numérique des deux sexes est donc basée sur une ségrégation égale lors de la méiose. Chez les humains, le sexe est déterminé par la présence ou l'absence du chromosome Y. Chez les mouches du fruit au contraire, le sexe est déterminé par le nombre de chromosomes X par rapport au nombre d'autosomes. Cette différence est démontrée par le sexe des combinaisons aberrantes, comme on le voit dans le tableau 4-1.

Les gènes autosomiques sont transmis de manière simple comme la couleur des graines des pois ou du pelage des souris. Les descendants mâles et femelles présentent exactement les mêmes rapports phénotypiques pour les caractères contrôlés par des gènes autosomiques. Par exemple, lors d'un croisement $A/a \times A/a$, les descendants femelles et mâles présenteront tous un rapport 3 : 1 de $A^- : a/a$. Ce n'est pas le cas des gènes situés sur le chromosome X car dans certains croisements, il y a des rapports phénotypiques différents pour la descendance mâle et la descendance femelle. L'observation de rapports phénotypiques différents dans la descendance mâle et femelle est de fait un diagnostic expérimental de gènes situés sur les chromosomes sexuels. Le mode exact de transmission est déterminé par la position du gène sur les chromosomes sexuels (les différentes régions sont indiquées dans la figure 4-15).

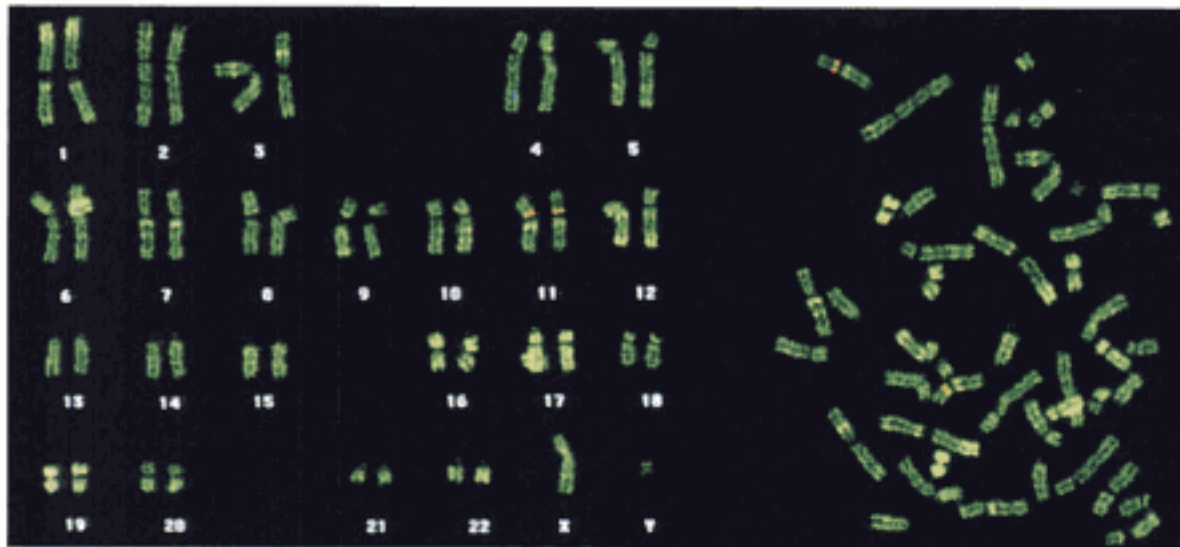


Figure 4-14 Chromosomes humains après coloration. Sous le microscope, les chromosomes ressemblent à un mélange désordonné, comme à droite. Cet ensemble est photographié. Ensuite, les différents chromosomes sont découpés de la photographie et regroupés en fonction de leur taille et de la disposition de leurs bandes, comme on le voit à gauche. La figure correspond aux chromosomes d'un homme.

La région spécifique au chromosome X n'a pas d'équivalent sur le Y. Elle contient de nombreux gènes avec des allèles produisant des phénotypes distincts. Ces gènes, pour la plupart, ne sont pas impliqués dans la détermination du sexe. Les gènes présents dans cette région sont appelés **gènes liés à l'X** et on dit que leur transmission est liée à l'X. Un cas typique de liaison à l'X est illustré par les allèles blanc et rouge d'un gène de la couleur de l'œil chez la drosophile (figure 4-16 à la page suivante). Remarquez les croisements dans lesquels les descendance mâle et femelle présentent des rapports phénotypiques différents. Par exemple, dans la F_2 du premier

croisement, il y a un rapport 3 : 1 entre le rouge et le blanc mais au contraire des cas de transmission des gènes autosomiques, les mouches blanches sont toutes des mâles. De plus, dans le second croisement, la descendance femelle dans la F_1 est entièrement rouge mais tous les mâles sont blancs. De tels modes de transmission constituent la marque distinctive permettant de déduire une liaison au chromosome X.

La région d'homologie (d'appariement) contient peu de gènes. Nous ne les aborderons donc pas plus en détail maintenant si ce n'est pour dire que dans chaque méiocyte, il y a dans cette région des événements

TABLEAU 4-1

Détermination chromosomique du sexe chez la drosophile et chez l'homme

Espèce	Chromosomes sexuels			
	XX	XY	XXY	XO
Drosophile	♀	♂	♀	♂
Homme	♀	♂	♂	♀

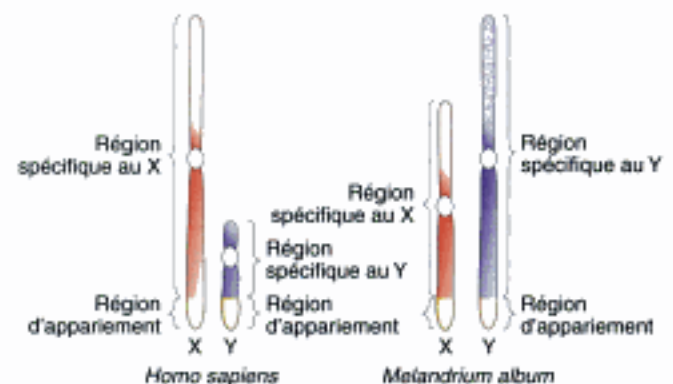


Figure 4-15 Région spécifique et région d'appariement des chromosomes sexuels de l'homme et de la plante *Melandrium album*. Les régions ont été déterminées en observant les endroits où les chromosomes s'appariaient au cours de la méiose et ceux où ils ne s'appariaient pas.

Hidden page

L'ANALYSE DES ARBRES GÉNÉALOGIQUES HUMAINS

On ne peut pas faire de croisements contrôlés dans l'espèce humaine. Les généticiens doivent donc se contenter d'examiner les archives d'une famille dans l'espoir d'y trouver des unions informatives, qui pourraient permettre de déduire la dominance et de distinguer la transmission autosomique de la liaison à l'X. L'enquêteur retrace l'histoire de certains phénotypes variants en remontant dans l'histoire familiale et dresse un **arbre généalogique** en utilisant les symboles indiqués dans la figure 4-17. Les indications contenues dans l'arbre généalogique doivent être interprétées différemment selon le cas où l'un des phénotypes distincts est une maladie rare ou bien si les deux phénotypes sont des morphes courants d'un même polymorphisme. Les maladies génétiques humaines peuvent présenter des phénotypes dominants ou récessifs et peuvent être soit autosomiques, soit liées à l'X. Ces quatre catégories seront traitées dans les paragraphes qui suivent.

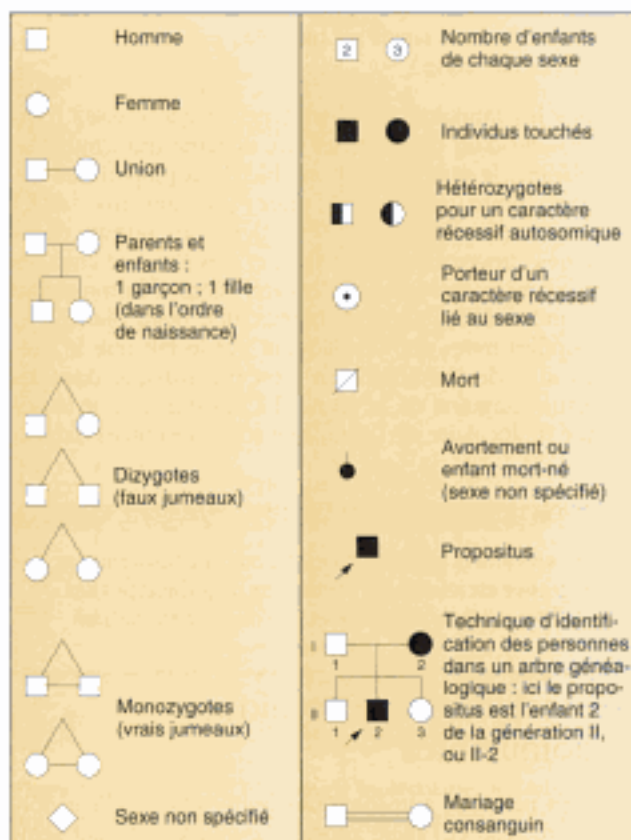
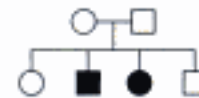


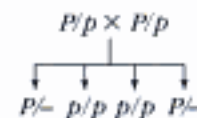
Figure 4-17 Symboles utilisés dans l'analyse d'arbres généalogiques humains. (D'après W. F. Bodmer et L. L. Cavalli-Sforza, *Genetics, evolution and man*. Copyright © 1976 par W. H. Freeman and Company.)

Les maladies autosomiques récessives

Le phénotype inhabituel d'une maladie récessive est déterminé par l'homozygotie pour un allèle récessif et le phénotype non touché est déterminé par l'allèle dominant correspondant. Nous avons vu au chapitre 3 que la phénylcétonurie (PCU) correspond à un phénotype récessif. La PCU est déterminée par un allèle que l'on peut appeler p et l'état normal par un allèle que l'on peut appeler P . Les personnes affectées par la maladie sont donc de génotype p/p et les personnes non touchées P/P ou P/p . Qu'observerait-on dans un arbre généalogique qui pourra nous révéler ce type de transmission ? Il y a deux éléments clés : la maladie apparaît généralement dans la descendance de parents non touchés et la descendance touchée comporte un nombre égal de filles et de garçons. Lorsque nous savons que la répartition phénotypique est identique chez les garçons et les filles, nous pouvons déduire que nous avons affaire à une transmission autosomique, et non à une liaison à l'X. L'arbre généalogique type qui suit illustre le fait que les enfants touchés naissent de parents sains.



Nous pouvons déduire immédiatement de ce schéma une transmission autosomique, avec l'allèle récessif responsable du phénotype exceptionnel (indiqué par la couleur noire). De plus, nous pouvons déduire que les deux parents doivent être hétérozygotes, P/p . (Tous deux doivent avoir un allèle p , car chacun en donne un à l'enfant touché et tous deux doivent également avoir un allèle P , car leur phénotype est normal.) Nous pouvons identifier les génotypes des enfants (dans l'ordre représenté) par $P/-$, p/p , p/p et $P/-$. L'arbre généalogique peut donc être réécrit de la façon suivante :



Remarquez une autre caractéristique intéressante de l'analyse des arbres généalogiques : même si les lois de Mendel s'appliquent, les rapports mendéliens sont rarement observés dans des familles considérées individuellement, car la taille des échantillons est trop faible. Dans l'exemple ci-dessus, nous observons un rapport phénotypique 1 : 1 dans la descendance de ce qui apparaît clairement comme un croisement monohybride pour lequel nous nous attendrions à un rapport 3 : 1. Si le couple avait, disons, 20 enfants, le rapport serait assurément voisin de 15 enfants sains et de 5 affectés de la PCU (le rapport 3 : 1 attendu pour un croisement monohybride) mais dans un échantillon de quatre enfants, tous les rapports sont possibles et d'ailleurs couramment observés.

Dans une population, les allèles récessifs rares sont présents généralement chez des hétérozygotes et non chez des homozygotes. Ceci est une simple question de probabilité : pour concevoir un homozygote récessif, les deux parents doivent posséder l'allèle p , alors que, pour concevoir un hétérozygote, il suffit que l'un des parents porte l'allèle. La formation d'un individu sain dépend généralement de l'union aléatoire d'hétérozygotes non apparentés et, pour cette raison, les arbres généalogiques correspondants apparaissent pauvres en phénotypes récessifs autosomiques et généralement seuls les frères et sœurs d'une même union sont affectés par la maladie.

Le mariage consanguin (mariage entre des personnes apparentées) augmente la probabilité d'une union entre deux hétérozygotes. La figure 4-18 présente l'exemple d'un mariage entre cousins. Les individus III-5 et III-6 sont cousins germains et ont deux enfants. On peut voir sur la figure qu'un ancêtre hétérozygote peut engendrer de nombreux descendants eux-mêmes hétérozygotes. Les mariages consanguins entraînent donc un risque plus élevé de donner naissance à des récessifs homozygotes.



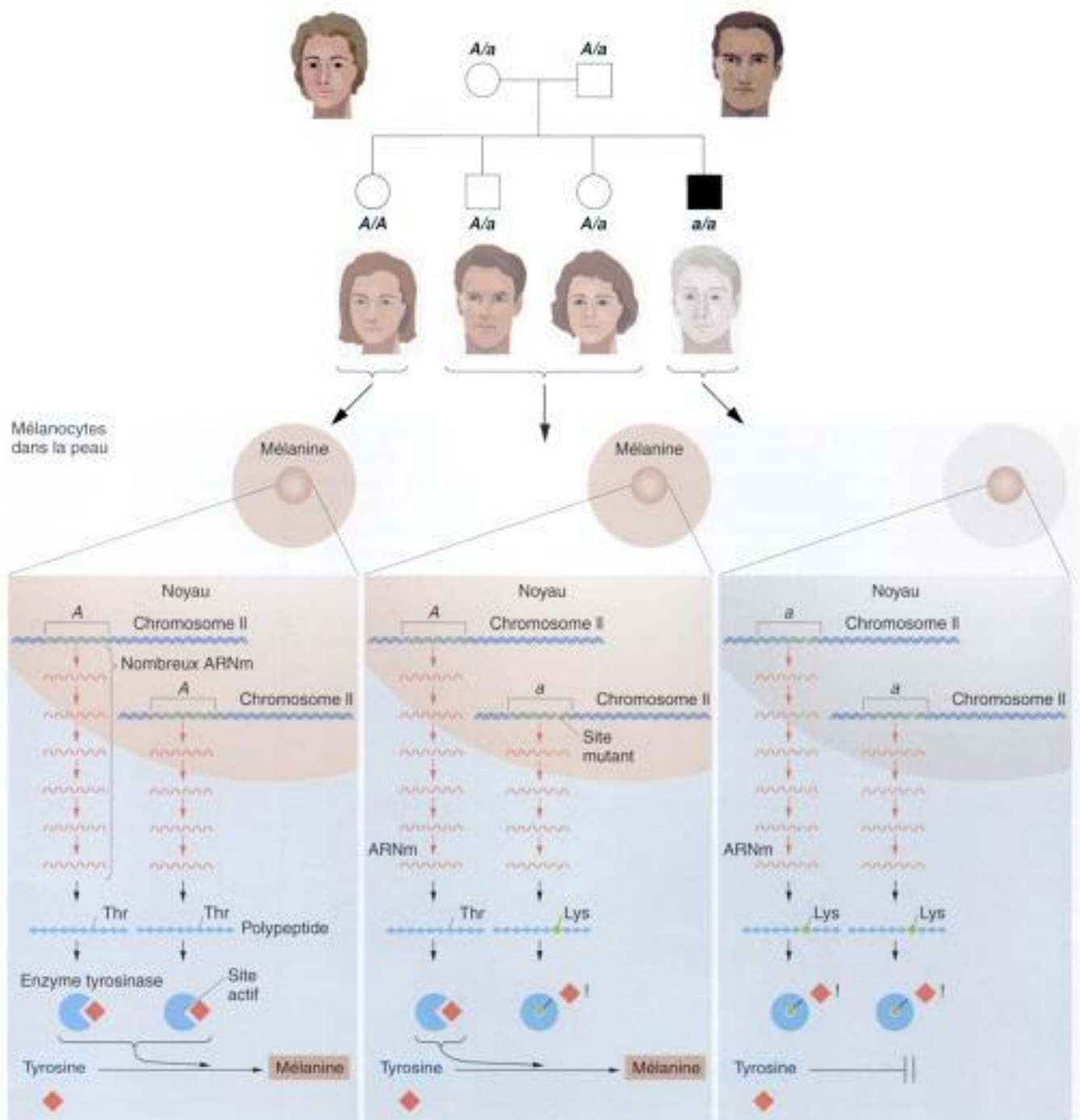


Figure 4-20 Génétique et biologie moléculaire de l'albinisme. Dans l'arbre généalogique, les parents hétérozygotes pour l'allèle récessif de l'albinisme engendrent trois enfants A/a , dont les cellules contiennent de la mélanine et un garçon a/a qui est albinos. Les trois schémas au bas de la figure montrent pourquoi il en est ainsi. Le gène de l'allèle A code un polypeptide nécessaire pour convertir la tyrosine en mélanine. Une mutation dans le gène de l'allèle a entraîne le remplacement d'une thréonine par une lysine. Le polypeptide résultant est incapable de catalyser la production de mélanine. L'allèle A chez les hétérozygotes code à lui seul suffisamment de polypeptide fonctionnel pour fabriquer de la mélanine. L'individu a/a ne produit pas de peptide fonctionnel et donc pas de mélanine.

forme de nanisme, est un exemple de phénotype dominant autosomique rare (voir figure 4-21). Dans ce cas, les gens de stature normale sont génotypiquement n/n et le phénotype nain devrait en principe être N/n ou N/N . On suppose cependant que chez les individus N/N , les deux « doses » de l'allèle D ont des conséquences si graves que ce génotype est létal. Si cela est vrai, tous les achondroplasiques sont hétérozygotes.

Dans l'analyse d'un arbre généalogique, les principaux éléments permettant d'identifier une maladie autosomique dominante sont la tendance du phénotype à se manifester à chaque génération de l'arbre généalogique et la transmission du phénotype des pères et des mères atteints, à leurs fils et à leurs filles. Une fois encore, la représentation des deux sexes parmi les descendants affectés est un argument contre une transmission liée à l' X . Le phénotype apparaît à chaque génération car généralement, l'allèle anormal porté par un individu provient d'un parent de la génération antérieure. (Des allèles anormaux peuvent apparaître *de novo* par mutation. C'est relativement rare mais il faut cependant garder cette possibilité à l'esprit.) Un arbre généalogique caractéristique d'une maladie dominante est présenté dans la figure 4-22. Une fois encore, remarquez que l'on n'observe pas forcément les rapports mendéliens dans une famille. Comme pour les maladies récessives, les individus portant une copie de l'allèle rare (A/a) sont bien plus courants que ceux qui portent les deux copies (A/A). La plupart des personnes touchées sont donc hétérozygotes et quasiment toutes les unions impliquant des maladies dominantes sont $A/a \times a/a$. Lorsque l'on totalise l'ensemble de la descendance de telles unions, on s'attend donc à un rapport 1 : 1 entre les individus non affectés (a/a) et les individus affectés (A/a).

La maladie de Huntington (appelée aussi « danse de Saint-Guy ») est un exemple de maladie autosomique

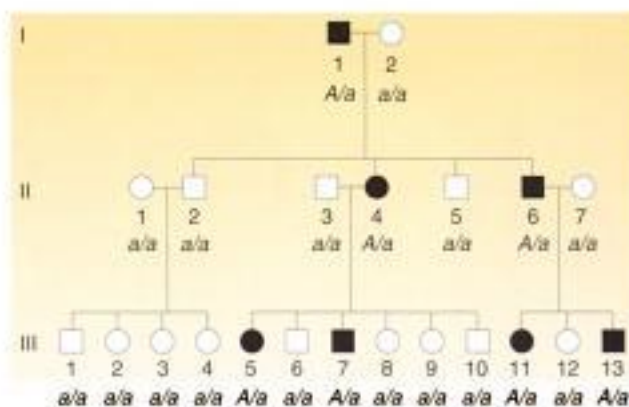


Figure 4-22 Arbre généalogique d'un phénotype dominant déterminé par un allèle dominant A . Dans cet arbre, tous les génotypes ont été déterminés par déduction.

dominante. Le phénotype correspond à une dégénérescence nerveuse, conduisant à des convulsions et à une mort prématurée. Toutefois, cette maladie ne s'exprime que tardivement et les symptômes n'apparaissent généralement qu'après que la personne a commencé à avoir des enfants. Chaque enfant d'un porteur de l'allèle anormal a une probabilité de 50 pour cent d'hériter de l'allèle anormal et de la maladie associée. Ce tragique état de fait a accéléré la recherche de moyens permettant d'identifier les personnes portant l'allèle anormal avant l'apparition de la maladie. La découverte de la nature moléculaire de l'allèle mutant et de mutations neutres dans l'ADN qui servent de « marqueurs » car elles sont situées à proximité de l'allèle touché sur le chromosome, a révolutionné ce type de diagnostic.



Figure 4-21 Le phénotype de l'achondroplasie humaine, illustré par une famille de cinq sœurs et deux frères. Le phénotype est déterminé par un allèle dominant, que nous pouvons appeler N , qui interfère avec la croissance des os pendant le développement. La plupart des membres de la population humaine peuvent être représentés par n/n pour ce gène. Cette photographie a été prise lors de l'arrivée de la famille en Israël après la fin de la seconde guerre mondiale. (UPI/Bettmann News Photos).

MESSAGE

Les arbres généalogiques de maladies autosomiques dominantes montrent des hommes et des femmes affectés à chaque génération et qui transmettent la maladie en proportions égales à leurs fils et à leurs filles.

Les populations humaines présentent de nombreux exemples de polymorphismes (généralement des dimorphismes) dans lesquels les phénotypes alternatifs du caractère sont déterminés par des allèles d'un même gène, par exemple les dimorphismes des fossettes au menton ou de leur absence, des lobes des oreilles soudés ou non, de l'implantation des cheveux en V sur le front ou non et ainsi de suite. L'interprétation des arbres généalogiques pour les dimorphismes est quelque peu différente de celle des maladies rares, car par définition, les morphes d'un dimorphisme sont courants. Regardons un arbre généalogique pour un dimorphisme humain intéressant. La plupart des populations humaines sont dimorphiques pour la capacité de percevoir le goût d'une substance chimique, le phénylthiocarbamide (PTC) : les gens peuvent soit détecter un goût amer très désagréable soit — à la grande surprise et à l'incrédulité des goûteurs — ne peuvent aucunement en sentir le goût. D'après l'arbre généalogique de la figure 4-23, on peut voir que deux goûteurs peuvent parfois engendrer des enfants non goûteurs. Ceci montre clairement que l'allèle de détection du goût est dominant et que l'allèle de l'incapacité correspondante est récessif. Remarquez cependant que presque toutes les personnes mariées à des membres de cette famille portaient l'allèle récessif, soit sous forme hétérozygote, soit sous forme homozygote. Un tel arbre diffère de ceux des maladies récessives rares pour lesquelles par convention, on suppose que toutes les personnes mariées à des membres de la famille sont des homozygotes normaux.

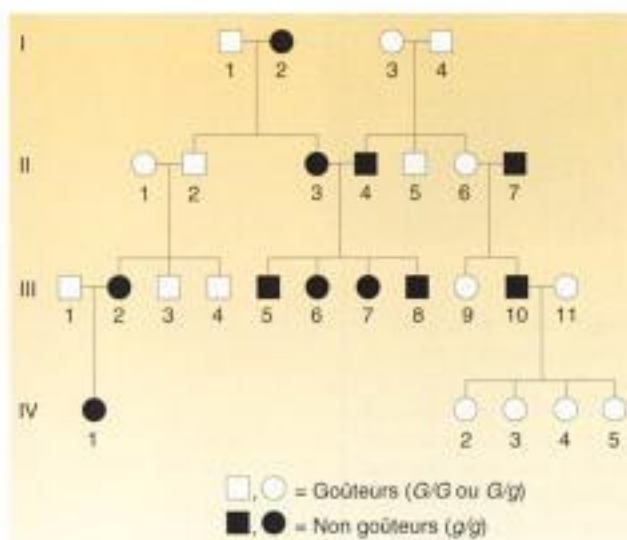


Figure 4-23 Arbre généalogique de la capacité à percevoir le goût de la substance PTC.

MESSAGE

Dans un polymorphisme, les morphes distincts sont souvent déterminés par les allèles d'un même gène autosomique.

Les maladies récessives liées à l'X

Peu de phénotypes déterminés par les allèles situés dans la région spécifique au chromosome X se rapportent à la détermination sexuelle. Les phénotypes dont la transmission est récessive et liée à l'X présentent typiquement les caractéristiques suivantes dans les arbres généalogiques :

1. Beaucoup plus d'hommes que de femmes présentent le phénotype étudié. Ceci est dû au fait qu'une femme présentant le phénotype ne peut être issue que d'une union entre une mère et un père portant tous deux l'allèle (par exemple, $X^a/X^a \times X^a/Y$), tandis qu'un homme présentant le phénotype peut apparaître lorsque la mère seule porte l'allèle. Si l'allèle récessif est particulièrement rare, presque tous les individus présentant le phénotype sont des hommes.
2. Aucun des descendants d'un homme affecté n'est lui-même atteint, mais toutes ses filles seront des « porteuses » hétérozygotes car les filles reçoivent l'un de leurs chromosomes X de leur père. La moitié des fils nés de ces filles porteuses seront affectés (Figure 4-24).

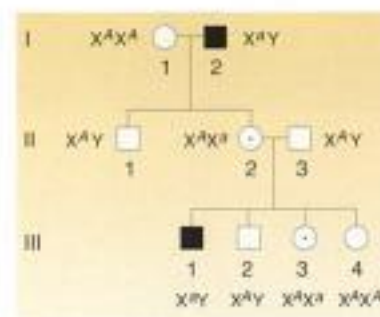
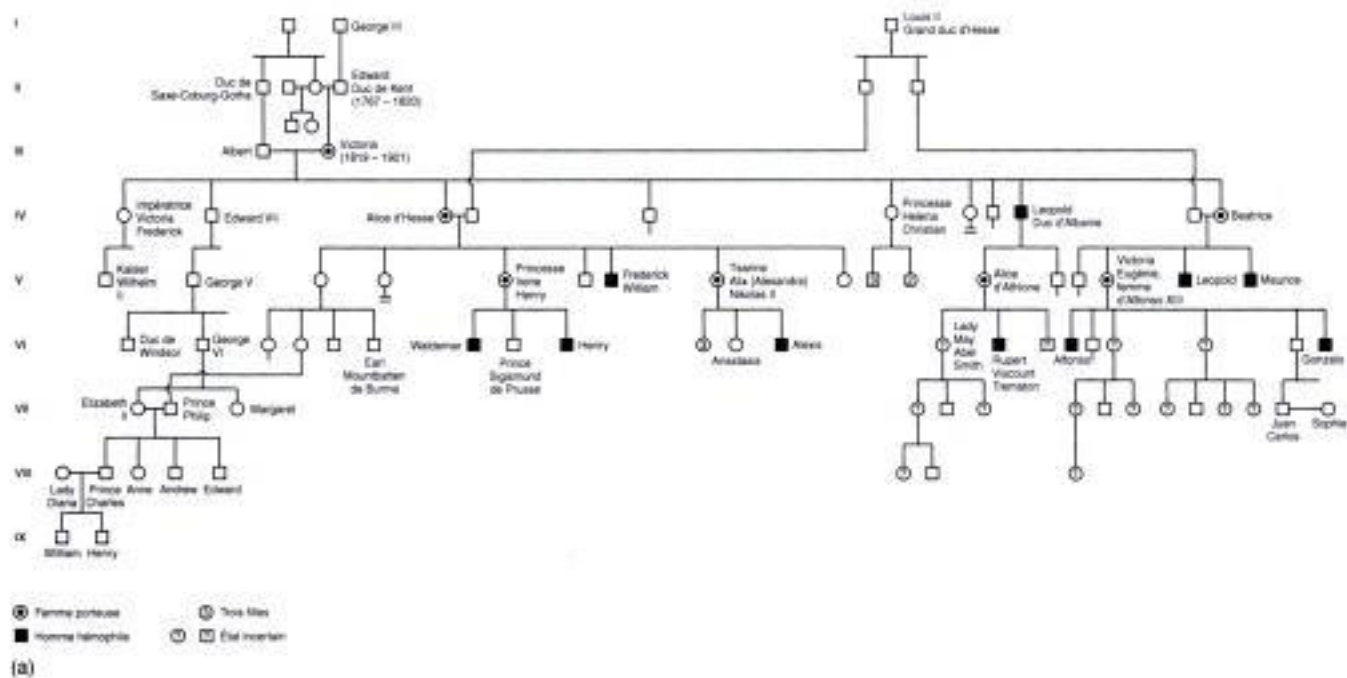


Figure 4-24 Arbre généalogique montrant que les allèles récessifs liés à l'X, exprimés chez les hommes sont ensuite transmis, sans être exprimés, par leurs filles pour être parfois exprimés à nouveau chez les fils de celles-ci. Remarquez que l'on ne peut distinguer phénotypiquement III-3 et III-4.



L'exemple sans doute le mieux connu est celui de l'hémophilie. Le sang des personnes affectées ne coagule pas. De nombreuses protéines doivent intervenir dans un ordre défini pour faire coaguler le sang. Le type le plus courant d'hémophilie est provoqué par l'absence ou le dysfonctionnement de l'une de ces protéines, appelée *facteur VIII*. Les cas d'hémophilie les plus célèbres se trouvent dans l'arbre généalogique des familles royales d'Europe (figure 4-25). L'allèle initial de l'hémophilie dans l'arbre généalogique est apparu spontanément (sous la forme d'une mutation) dans les cellules germinales des parents de la reine Victoria ou de la reine Victoria elle-même. Alexis, le fils du dernier tsar de Russie reçut l'allèle de la reine Victoria qui était la grand-mère de sa mère Alexandra. De nos jours, l'hémophilie peut être soignée, mais autrefois il s'agissait d'une maladie potentiellement mortelle. Il est intéressant de remarquer que dans le Talmud juif il existe des règles sur l'exemption de la circoncision pour certains garçons, qui montrent clairement que le mode de transmission de la maladie par des femmes porteuses non touchées était bien compris dans les temps anciens. Par exemple, les fils des femmes dont les sœurs avaient eu des fils qui avaient saigné abondamment lors de la circoncision, étaient dispensés.

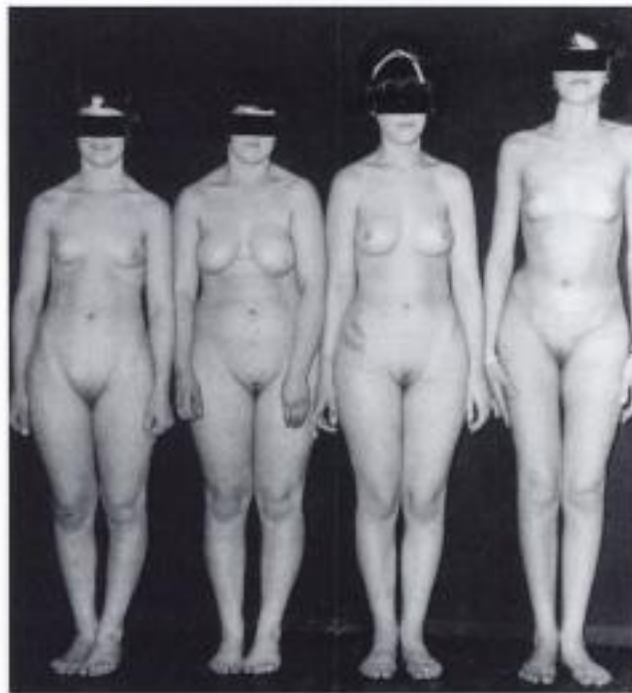


Figure 4-26 Quatre frères présentant un syndrome de féminisation testiculaire (insensibilité congénitale aux androgènes). Les quatre sujets de cette photographie possèdent 44 autosomes ainsi qu'un X et un Y mais ils ont reçu l'allèle récessif lié à l'X responsable de l'insensibilité aux androgènes (hormones mâles). Une de leurs sœurs (non présentée) qui était génétiquement XX était porteuse de l'allèle et eut un enfant qui présentait également un syndrome de féminisation testiculaire. (Leonard Pinsky, McGill University).

La dystrophie musculaire de Duchenne est une maladie récessive fatale liée à l'X. Le phénotype est une dégénérescence et une atrophie des muscles. La maladie se manifeste généralement avant l'âge de 6 ans. L'enfant doit utiliser un fauteuil roulant vers 12 ans et meurt vers 20 ans. Le gène de la dystrophie musculaire de Duchenne a été isolé et on sait qu'il code une protéine du muscle, la dystrophine. Une telle avancée permet d'espérer une meilleure compréhension de la physiologie de cette maladie et pour finir, de concevoir une thérapie.

Un phénotype récessif rare lié à l'X qui est intéressant du point de vue de la différenciation sexuelle est une maladie appelée *syndrome de féminisation testiculaire* dont la fréquence est d'environ 1 sur 65 000 naissances de garçons. Les personnes affectées par ce syndrome sont chromosomiquement des hommes, 44A + XY mais elles se développent comme des femmes (figure 4-26). Elles ont des organes génitaux externes féminins, un vagin et pas d'utérus. Les testicules peuvent être présents soit dans les lèvres vulvaires soit dans l'abdomen. Bien qu'un grand nombre de ces personnes soient mariées avec bonheur, elles sont bien sûr stériles. Leur état ne peut être inversé par un traitement par une hormone mâle (androgène). C'est pourquoi on appelle parfois cette maladie *syndrome d'insensibilité aux androgènes*. Cette insensibilité s'explique par le fait que l'allèle responsable code une protéine qui sert de récepteur aux androgènes et qui fonctionne mal. L'hormone mâle ne peut donc avoir d'effet sur les organes cibles impliqués dans la masculinité. Chez les humains, les caractères féminins apparaissent lorsque le système déterminant les caractères masculins n'est pas fonctionnel.

Les maladies dominantes liées à l'X

Les arbres généalogiques des rares phénotypes dominants liés à l'X présentent les caractéristiques suivantes :

1. Les hommes affectés transmettent la maladie à toutes leurs filles mais à aucun de leurs fils (figure 4-27).
2. Les femmes mariées à des hommes non affectés transmettent la maladie à la moitié de leurs fils et de leurs filles.

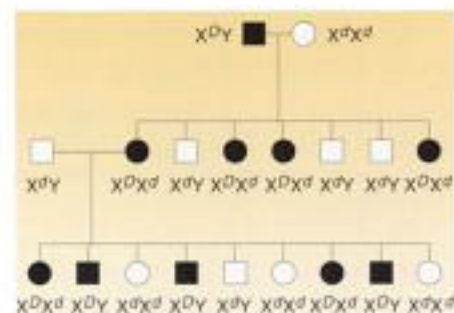


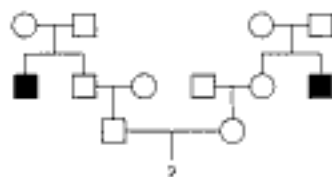
Figure 4-27 Arbre généalogique d'une maladie dominante liée à l'X.

Il existe peu d'exemples de phénotypes dominants liés à l'X chez l'homme. L'un est l'hypophosphatémie, une sorte de rachitisme résistant à la vitamine D.

Les mécanismes de dominance et de récessivité liées à l'X chez l'homme sont quelque peu compliqués par le phénomène d'inactivation d'un chromosome X qui se produit chez les mammifères. Ce sujet sera traité au chapitre 16.

Calculer les risques lors de l'analyse d'arbres généalogiques

Lorsque l'on sait qu'un allèle de maladie est présent dans une famille, on peut utiliser la connaissance des modalités simples de transmission des gènes pour calculer la probabilité que des futurs parents aient un enfant atteint de cette maladie. Par exemple, un couple marié découvre que chacun a un oncle affecté de la maladie de Tay-Sachs (une maladie autosomique récessive grave). L'arbre généalogique se présente ainsi :



La probabilité qu'ils aient un enfant affecté de la maladie de Tay-Sachs peut être calculée de la façon suivante. Il faut déterminer si l'homme et la femme sont hétérozygotes (on sait qu'ils n'ont pas cette maladie) car si tous deux sont hétérozygotes, alors ils risquent d'avoir un enfant malade.

1. Les grands-parents de l'homme devaient tous deux être hétérozygotes T/t car ils ont donné naissance à un enfant t/t (l'oncle). Il s'agit donc bien d'une union monohybride. Le père de l'homme concerné pourrait être T/T ou T/t mais on sait que les probabilités relatives de ces génotypes sont de $1/4$ et $1/2$ respectivement (le rapport attendu dans la descendance d'un croisement monohybride est de $1/4 T/T$, $1/2 T/t$ et $1/4 t/t$). Il y a donc une probabilité de $2/3$ que le père soit un hétérozygote [calculé de la façon suivante : $1/2$ divisé par $(1/4 + 1/2)$].
2. On doit supposer que la mère de l'homme est T/t car elle est arrivée par mariage dans la famille et les allèles de la maladie sont généralement rares. Donc si le père est T/t , alors son union avec la mère a été un croisement $T/t \times T/t$ et les proportions attendues des descendants sont $1/2 T/T$ et $1/2 T/t$.
3. La probabilité totale que l'homme soit hétérozygote doit être calculée en utilisant une règle statistique

appelée **règle du produit**, où il est dit que la *probabilité que deux événements indépendants se produisent tous les deux est le produit de leurs probabilités respectives*. La probabilité que l'homme soit hétérozygote est la probabilité que le père soit hétérozygote multipliée par la probabilité que le père ait un fils hétérozygote, soit $2/3 \times 1/2 = 1/3$.

4. De la même façon, la probabilité que l'épouse de cet homme soit hétérozygote est de $1/3$.
5. S'ils sont tous deux hétérozygotes (T/t), alors la probabilité qu'ils aient un enfant t/t est de $1/4$. La probabilité totale que le couple ait un enfant atteint est de $1/3 \times 1/3 \times 1/4 = 1/36$; en d'autres termes, une « chance » sur 36.

LA TRANSMISSION DES GÈNES D'ORGANITES

Les mitochondries et les chloroplastes sont des organites spécialisés, localisés dans le cytoplasme. Ils contiennent une partie bien identifiée du génome de la cellule (les gènes mitochondriaux sont présentés dans la figure 4-28). Ces gènes ont leur propre mode de transmission. Dans un croisement, les deux parents contribuent de façon égale au génome nucléaire du zygote. Cependant, la contribution cytoplasmique du père et de la mère est généralement inégale ; l'ovule fournit quasiment tout le cytoplasme et le spermatozoïde presque rien. Les organites résident dans le cytoplasme, c'est pourquoi les gènes des organites sont généralement **transmis exclusivement par la mère**. En d'autres termes, quasiment aucun ADN d'organe du zygote ne provient du père. On voit un exemple simple chez le mutant *poky* (littéralement *lambin*) à croissance ralentie de *Neurospora*, qui apparaît à la suite d'un défaut dans un des gènes mitochondriaux. On peut croiser *Neurospora* de telle façon qu'un parent joue le rôle de parent maternel et l'autre, de parent paternel. Dans le croisement d'un type femelle *poky* avec un type mâle normal, tous les descendants sont *poky* (figure 4-29), ce qui est exactement le résultat attendu pour la transmission d'un gène mitochondrial.

La figure 4-30 de la page 120 démontre la transmission maternelle chez la plante belle-de-nuit. La couleur des chloroplastes de cette plante détermine la couleur des différentes branches. Les branches panachées sont des mosaïques de cellules toutes blanches ou toutes vertes. Les fleurs peuvent provenir de branches vertes, blanches ou panachées mais lorsqu'on les croise, c'est l'ovule qui détermine la couleur de la branche dans la plante résultante. Par exemple, si l'ovule provient d'une fleur de branche blanche, quelle que soit la provenance du pollen, la plante résultante aura des branches blanches, ce

Figure 4-28 Cartes des ADNmt de levure et d'homme. Chaque carte est présentée sous la forme de deux cercles concentriques correspondant aux deux brins de l'hélice d'ADN. Remarquez que les mutants utilisés dans l'analyse d'ADNmt de levure sont indiqués à l'opposé des gènes structuraux correspondants. Vert = exons et gènes non morcelés, rouge = gènes d'ARNt et jaune = URF (*unassigned reading frames*; cadres de lecture non attribués). Les gènes d'ARNt sont indiqués par l'abréviation de leurs acides aminés. Les gènes ND codent des sous-unités de la NADH déshydrogénase. (Remarquez que la carte humaine n'est pas dessinée à la même échelle que la carte de levure)

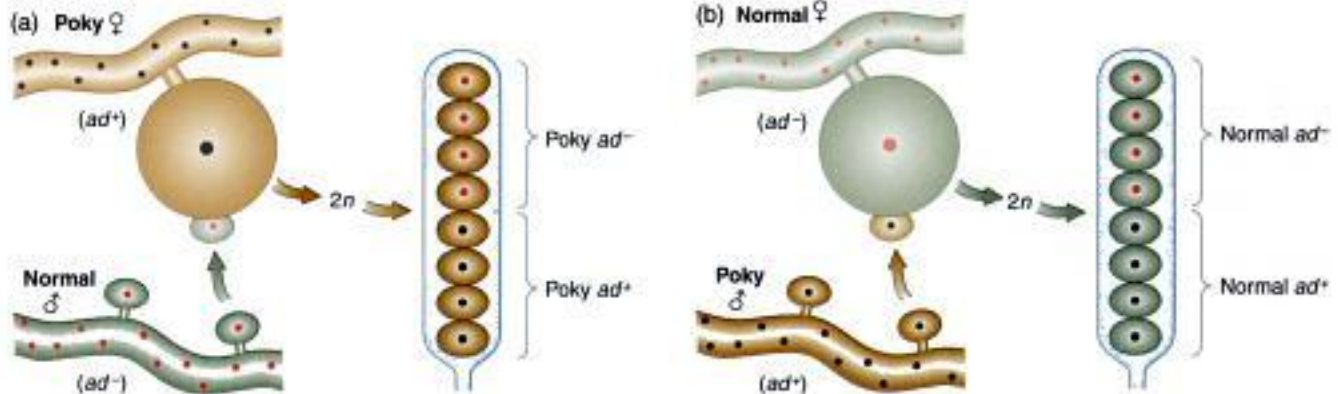
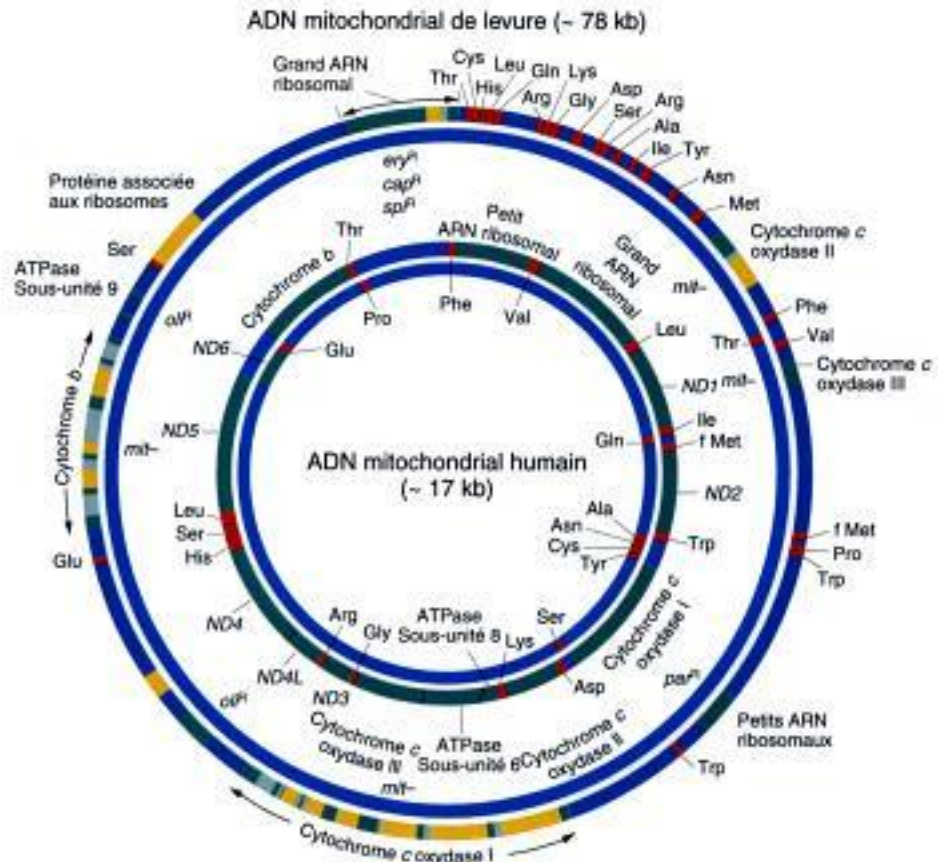


Figure 4-29 Explication des différents résultats des croisements réciproques réalisés entre des souches de *Neurospora poky* et normales. Le parent contribuant à la majeure partie du cytoplasme des cellules de la descendance est appelé *féfemelle*. L'ombrage en marron représente le cytoplasme avec les déterminants *poky*. Le locus nucléaire avec les allèles *ad⁻* et *ad⁺* est utilisé pour illustrer la ségrégation des gènes nucléaires dans le rapport mendélien 1 : 1 attendu.

(a) Plante panachée

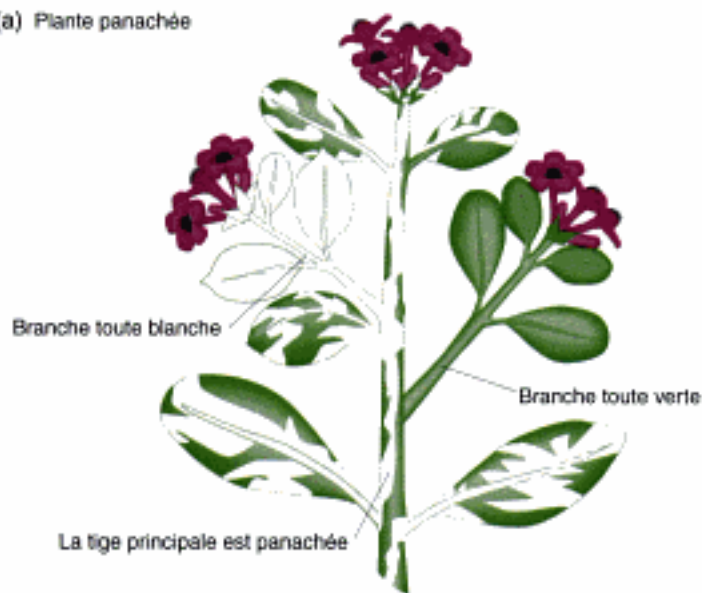


Figure 4-30 Panachure des feuilles chez la plante *Mirabilis jalapa*, la belle-de-nuit. (a) Des fleurs peuvent se former sur n'importe quelle branche (panachée, verte ou blanche) et ces fleurs peuvent être utilisées dans des croisements. (b) Dans les croisements entre fleurs provenant de branches de couleurs différentes, la couleur du zygote et de la plante résultante est déterminée par transmission maternelle. Les deux premiers croisements montrent qu'il s'agit d'une transmission maternelle exclusive. Si la branche maternelle est panachée, trois types de zygotes peuvent apparaître, selon que l'ovule contient exclusivement des chloroplastes verts, ou des chloroplastes blancs ou possède les deux sortes. La transmission maternelle s'applique encore aux trois types mais dans les zygotes panachés, un processus de ségrégation cytoplasmique au cours des divisions cellulaires qui suivent produit une mosaïque de cellules soit toutes blanches soit toutes vertes, et donc une plante panachée.

(b) Résultats des croisements entre branches

Oosphère (n)	Cellule pollinique du mâle (n)	Constitution du zygote ($2n$)
Blanche ♀	N'importe quel ♂	Blanche
Verte ♀	N'importe quel ♂	Verte
Panachée ♀	N'importe quel ♂	Blanche
Oosphère de type 1		Blanche
Oosphère de type 2		Verte
Oosphère de type 3		<div> <div>Division cellulaire</div> <div> </div> </div>

Panachées

qui prouve une transmission maternelle. Les zygotes panachés (bas de la figure 4-30) sont des mélanges cytoplasmiques de deux types de chloroplastes et ce qui est intéressant, ces deux types subissent une ségrégation cytoplasmique au cours de la division cellulaire, conduisant à des secteurs distincts vert et blanc responsables de la panachure (bigarrure) dans les branches.

MESSAGE

Les phénotypes variants provoqués par des mutations de l'ADN d'organite sont généralement transmis par la mère.

RÉSUMÉ

L'élément central de la transmission des gènes au travers des générations de cellules ou d'organismes est la réplication de l'ADN. La réplication de l'ADN précède à la fois la division cellulaire sexuée et asexuée. Au cours de la réplication de l'ADN, les deux moitiés de la double hélice se séparent et les simples-brins servent de matrice pour une nouvelle polymérisation des nucléotides. Chaque molécule fille d'ADN est donc à moitié ancienne et à moitié nouvelle ; ce mécanisme de réplication est dit semi-conservatif. L'ADN polymérase catalyse la nouvelle polymérisation en ajoutant des nucléotides seulement aux extrémités 3'. Les deux matrices étant antiparallèles, la réplication en 3' sur une matrice se fait par une polymérisation continue (il s'agit du brin précoce), mais sur l'autre matrice, la synthèse se fait de façon discontinue en sens opposé, produisant des fragments d'Okazaki, qui sont ensuite reliés les uns aux autres (cette chaîne est le brin retardé).

La division cellulaire asexuée s'accompagne d'un type de division nucléaire appelé *mitose*, qui répartit les chromosomes de telle sorte que les deux cellules filles contiennent exactement les mêmes gènes que la cellule d'origine. La division cellulaire sexuée se déroule dans des cellules diploïdes spécialisées appelées *métiocytes*, qui se divisent deux fois pour produire quatre cellules haploïdes (qui deviennent les ovules et les spermatozoïdes chez les animaux et les plantes). La division cellulaire sexuée s'accompagne de deux divisions nucléaires constituant la *méiose*, qui répartissent les chromosomes de

façon à produire des produits de méiose possédant un seul jeu de chromosomes.

Les rapports phénotypiques dans la descendance des croisements sont le résultat de mouvements hautement programmés des chromosomes lors de la méiose. Les allèles d'une même paire, occupant une position sur une paire de chromosomes homologues se séparent et migrent dans des cellules sexuelles différentes car les chromosomes homologues s'apparient physiquement et subissent une ségrégation lors de la première division de méiose. Les produits de la méiose d'un hétérozygote A/a sont pour $1/2 A$ et pour $1/2 a$. À l'issue de l'autofécondation d'un hétérozygote, on observe dans la descendance un rapport 3 : 1 des phénotypes dominants sur les récessifs. Si l'on réalise un croisement-test avec a/a alors le rapport est de 1 : 1.

Les gènes situés dans la région spécifique au chromosome X peuvent présenter des modes de transmission dans lesquels les descendants mâle et femelle apparaissent dans des rapports phénotypiques différents.

Les modes de transmission chez l'homme peuvent être déduits de l'analyse d'arbres généalogiques. À l'inverse, on peut utiliser la connaissance des modes de transmission pour prédire le risque qu'un couple ait un enfant affecté d'une maladie spécifique qui existe dans la famille.

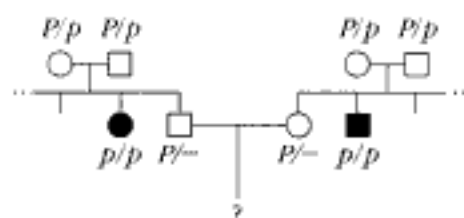
Les phénotypes et les gènes d'organites sont transmis par la mère car presque tout le cytoplasme des descendants provient de celle-ci.

DIAGRAMME CONCEPTUEL

Dessinez un diagramme conceptuel mettant en relation le maximum de termes parmi ceux qui sont proposés. Remarquez que ces termes ne sont pas écrits dans un ordre particulier.

méiose / chromosome / chromatide / réplication d'ADN / autofécondation / ségrégation égale / hétérozygote A/a / rapport 1 : 1 / rapport 3 : 1 / cellules sexuelles

Hidden page



La seule façon pour le couple d'avoir un enfant atteint de PCU est qu'ils soient tous deux hétérozygotes (il est évident qu'ils n'ont pas eux-mêmes la maladie). Les deux unions des grands-parents sont de simples croisements monohybrides ($P/p \times P/p$) dont la descendance doit se présenter dans les proportions indiquées :

$$\left. \begin{array}{l} \frac{1}{4} P/P \\ \frac{1}{2} P/p \\ \frac{1}{4} p/p \end{array} \right\} \begin{array}{l} \text{normal} \\ \\ \text{PCU} \end{array}$$

Nois savons que l'homme et la femme sont normaux, aussi pour chacun d'eux, la probabilité d'être un hétérozygote est de $2/3$ car à l'intérieur de la classe $P/-$ normale, $2/3$ sont P/p et $1/3$ P/P . En utilisant la règle du produit, la probabilité que l'homme et la femme soient tous deux hétérozygotes est donc de $2/3 \times 2/3 = 4/9$. S'ils sont tous deux hétérozygotes, alors un quart de leurs enfants auront la PCU. La probabilité que leur premier enfant ait la PCU est donc de $1/4$ et la probabilité qu'ils soient hétérozygotes et que leur premier enfant ait la PCU est de $4/9 \times 1/4 = 4/36 = 1/9$, ce qui répond à la question.

PROBLÈMES

- Supposez que vous possédiez des cellules eucaryotes diploïdes qui se diviseront de façon asexuée dans une solution nutritive particulière à l'intérieur d'un tube à essai. Votre point de départ est constitué de cellules qui ne se sont pas engagées dans la phase S prémitotique (phase de réplication). Vous placez les cellules dans une solution nutritive contenant des nucléotides radioactifs qui peuvent être incorporés dans de l'ADN néosynthétisé par l'ADN polymérase. En utilisant la couleur rouge pour le brin d'ADN radioactif et la couleur noire pour le brin non radioactif, dessinez la double hélice d'un chromosome dans les situations suivantes :
 - Au début de l'expérience (avant la phase S prémitotique)
 - Après la formation des chromatides
 - Avant la phase S prémitotique dans les cellules filles
 - Avant la phase S prémitotique dans leurs cellules filles (c'est-à-dire après deux divisions dans la solution de nucléotides radioactifs).
- Dessinez un graphique représentant le contenu en ADN en fonction du temps, dans une cellule diploïde qui passe d'abord par la mitose et ensuite par la méiose.

- Pourquoi la synthèse d'ADN est-elle continue sur une matrice et discontinue sur l'autre ?
- Considérons le segment d'ADN suivant, qui appartient à une molécule bien plus longue constituant un chromosome.

....5'-ATTCGTACGATCGACTGACTGACAGTC-3'.....

....3'-TAAGCATGCTAGCTGACTGACTGTCAG-5'.....

Si l'ADN polymérase commence à répliquer ce segment par la gauche,

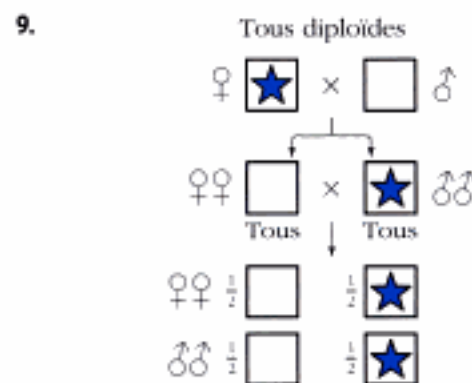
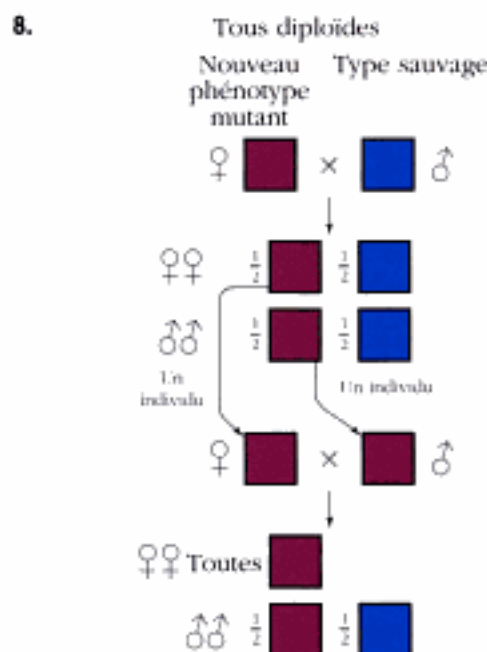
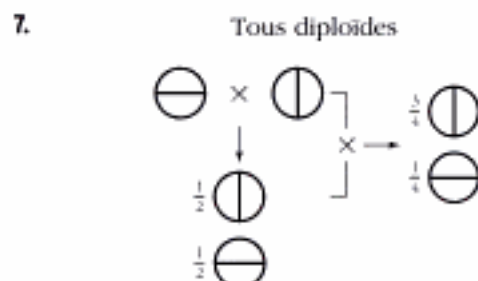
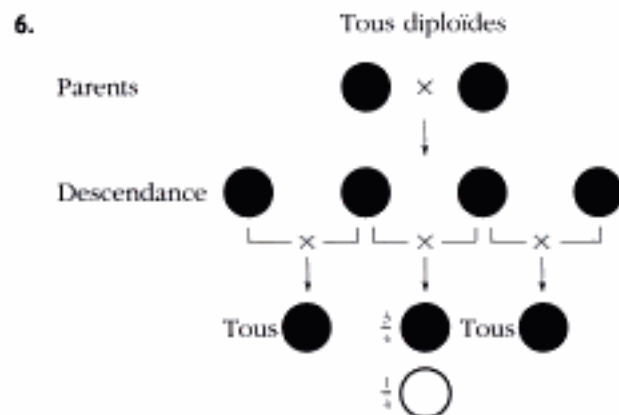
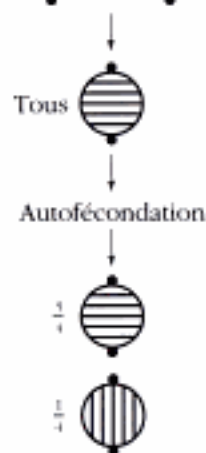
- Quelle sera la matrice du brin précoce ?
- Dessinez la molécule lorsque l'ADN polymérase est à mi-parcours de ce segment.
- Dessinez les deux molécules filles en entier.
- Votre dessin en b est-il compatible avec une réplication bidirectionnelle se déroulant à partir d'une seule origine, ce qui est le mode de réplication habituel ?

Problèmes de reconnaissance de formes

Dans les problèmes 5 à 9, les schémas symbolisent les phénotypes et les résultats d'analyses de croisements. Déduisez les génotypes des individus présentés dans chaque diagramme de façon aussi poussée que possible,

en utilisant des symboles d'allèles qui prennent en compte toutes les indications de dominance dans les données.

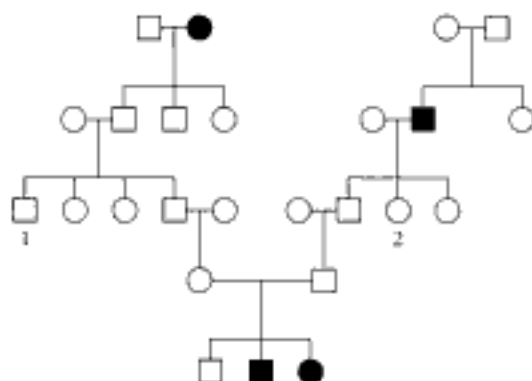
5. Diploïde de lignée pure  × Diploïde de lignée pure 



10. Si vous aviez une drosophile (*Drosophila melanogaster*) de phénotype A^- , quel test réaliseriez-vous pour déterminer si elle est A/A ou A/a ?
11. Deux cochons d'Inde noirs ont été croisés et pendant plusieurs années, ils ont produit 29 descendants noirs et 9 descendants blancs. Expliquez ces résultats en indiquant les génotypes des parents et des descendants.
12. Les taureaux de race frisonne sont normalement noir et blanc. Un superbe taureau noir et blanc, Charlie, a été acheté 100 000 \$ par un fermier. Tous les descendants de Charlie sont normaux en apparence. Cependant, certaines paires de ses descendants, lorsqu'on les croise, donnent des descendants rouge et blanc à une fréquence voisine de 25

Hidden page

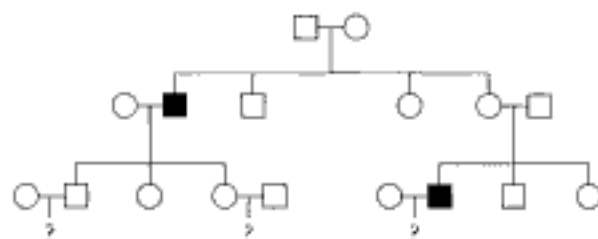
- a. La maladie est-elle transmise comme un phénotype dominant ou récessif ? Justifiez votre réponse.
 - b. Donnez les génotypes du maximum d'individus dans l'arbre généalogique. (Inventez vos propres symboles d'allèles et définissez-les).
 - c. Considérez les quatre enfants non touchés des parents III-4 et III-5. Parmi les descendants des quatre enfants issus des parents ayant ces génotypes, à votre avis quelle sera la proportion d'enfants non affectés ?
16. L'arbre généalogique ci-dessous concerne une maladie rare du rein.



- a. Déduisez le mode de transmission de cette maladie en précisant vos raisons.
 - b. Si les individus 1 et 2 se marient, quelle est la probabilité que leur premier enfant soit affecté de la maladie du rein ?
17. Un croisement est réalisé entre une levure haploïde auxotrophe pour la leucine (c'est-à-dire ayant besoin de leucine pour croître), de génotype *leu* et une souche de type sauvage capable de fabriquer sa propre leucine (*leu*⁺).
- a. Quel sera le génotype du méiocyte transitoire ?
 - b. Quels seront les génotypes des descendants et en quelles proportions apparaîtront-ils ?
 - c. Dessinez des schémas qui montrent les structures possibles des allèles au niveau de l'ADN ainsi que la nature de leurs produits protéiques.
18. L'allèle récessif *s* conduit à de petites ailes chez la drosophile et l'allèle *s*⁺ à des ailes normales. On sait que le gène est lié à l'X. (Note : Il ne s'agit pas

du même gène que celui traité dans ce chapitre). Si un mâle à petites ailes est croisé avec une femelle homozygote de type sauvage, quel rapport entre mouches à petites ailes et mouches à ailes normales attendez-vous dans chaque sexe de la F₁ ? Si les mouches de la F₁ sont croisées entre elles, quel rapport attendez-vous chez les descendants (F₂) ? Quel rapport peut-on prédire si l'on réalise un croisement en retour (*backcross*) entre les femelles F₁ et leur père ?

19. La dystrophie musculaire de Duchenne est liée au sexe et n'affecte généralement que les garçons. Les victimes de cette maladie s'affaiblissent progressivement et cela dès l'enfance.
- a. Quelle est la probabilité qu'une femme dont le frère est atteint de la maladie ait un enfant affecté ?
 - b. Si le frère de votre mère (votre oncle) est atteint de myopathie de Duchenne, quelle est la probabilité que vous ayez reçu l'allèle ?
 - c. Si le frère de votre père a la maladie, quelle est la probabilité que vous ayez reçu l'allèle ?
20. L'arbre généalogique suivant concerne une maladie rare incapacitante mais non mortelle.



- a. Déterminez le mode de transmission de cette maladie.
 - b. Écrivez le génotype de chaque individu d'après le mode de transmission que vous proposez.
 - c. Si vous étiez le médecin de cette famille que diriez-vous à chacun des trois couples de la troisième génération quant à la probabilité d'avoir un enfant touché ?
21. Un allèle mutant chez des souris leur donne une queue coudée. Six couples de souris ont été croisés. Leurs phénotypes et ceux de leur descendance sont indiqués ci-dessous. N est le phénotype normal ; C le phénotype coudé. Déduisez le mode de transmission de ce phénotype.

Croisement	Parents		Enfants	
	♀	♂	♀	♂
1	N	C	Tous C	Tous N
2	C	N	$\frac{1}{2}$ B, $\frac{1}{2}$ N	$\frac{1}{2}$ C, $\frac{1}{2}$ N
3	C	N	Tous C	Tous C
4	N	N	Tous N	Tous N
5	C	C	Tous C	Tous C
6	C	C	Tous C	$\frac{1}{2}$ C, $\frac{1}{2}$ N

- Ce phénotype est-il récessif ou dominant ?
- Est-il autosomique ou lié à l'X ?
- Indiquez les génotypes de tous les parents et de leurs descendants.

22. Une nouvelle souche mutante du champignon haploïde *Neurospora* a été découverte. Elle présente une croissance irrégulière, par à-coups, c'est pourquoi on la désigne par « stop » (*Neurospora* croît généralement de façon continue : souche « continue »). La souche stop a été croisée avec une autre souche mutante, qui est de couleur jaune au lieu de la couleur orange habituelle. Le croisement a été réalisé de manière réciproque en utilisant chaque souche soit comme parent paternel soit comme parent maternel. Les résultats ont été les suivants :

Croisement 1

Parents	type femelle stop, orange × type mâle continue, jaune
Descendants	$\frac{1}{2}$ stop, orange $\frac{1}{2}$ stop, jaune

Croisement 2

Parents	type femelle continue, jaune × type mâle stop, orange
Descendants	$\frac{1}{2}$ continue, jaune $\frac{1}{2}$ continue, orange

- Que pouvez-vous conclure quant à la localisation de ces gènes dans la cellule ?
- Expliquez clairement la différence entre les résultats des croisements réciproques.

23. Un homme dont le grand-père était atteint de galactosémie (une maladie autosomique rare) s'est marié avec une femme dont la sœur est atteinte de la même maladie. La femme est enceinte de leur premier enfant.

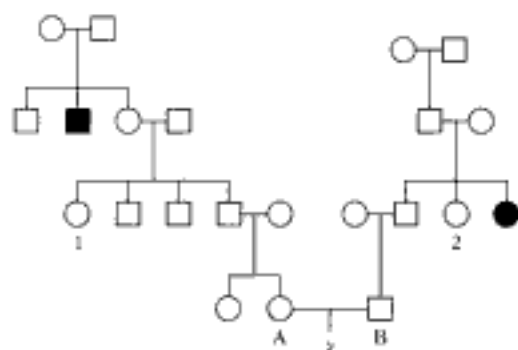
- Dessinez l'arbre généalogique décrit.

- Quelle est la probabilité que cet enfant soit atteint de galactosémie ?
- Si le premier enfant est effectivement atteint, quelle est la probabilité que le deuxième enfant soit lui aussi atteint ?

24. Un homme dont le grand-père maternel est atteint de favisme (une sensibilité à l'ingestion des fèves, récessive, rare et liée à l'X) se marie avec une femme dont l'oncle maternel est atteint de la même sensibilité.

- Dessinez l'arbre généalogique tel qu'il est décrit.
- Quelle est la probabilité que leur premier enfant soit atteint de favisme ?
- Si leur premier enfant est atteint de favisme, quelle est la probabilité que leur deuxième enfant en soit également atteint ?

25. L'arbre généalogique suivant concerne une maladie héréditaire rare du système nerveux :



- À votre avis, la maladie est-elle récessive ou dominante ? (Expliquez.)
 - La maladie est-elle autosomique ou liée à l'X ? (Expliquez.)
 - D'après vos réponses, quelle est la probabilité que le premier enfant du couple A et B soit atteint de la maladie ?
26. Chez la plante *Mimulus*, les tiges peuvent être soit lisses soit velues et cette différence est déterminée par une différence allélique simple. Examinez les croisements suivants et indiquez le génotype des plantes parentales et des plantes de la descendance dans chaque croisement. (Inventez vos propres symboles et expliquez leur signification.)

Croisement	Parents	Descendants
1	Plante 1 (velue) × plante 2 (velue)	Toutes velues
2	Plante 1 (velue) × plante 3 (velue)	$\frac{3}{4}$ velues $\frac{1}{4}$ lisses
3	Plante 1 (velue) × plante 4 (lisse)	$\frac{3}{4}$ velue $\frac{1}{4}$ lisse
4	Plante 2 (velue) × plante 4 (lisse)	Toutes velues



27. Représentez la descendance du croisement $A/a \times A/a$ à la fois sous la forme d'un tableau et sous la forme d'un arbre généalogique. Quels sont les avantages et les inconvénients de chaque représentation ?

28. Considérons les arbres généalogiques suivants qui correspondent à deux maladies rares différentes :



- Déduisez-en le mode de transmission le plus probable de ces deux maladies et expliquez votre raisonnement.
- Calculez les probabilités que le couple A et B et le couple C et D aient un enfant affecté.

5

LA RECOMBINAISON DES GÈNES



POUR L'ESSENTIEL

Deux gènes proches l'un de l'autre sur un même chromosome ne sont pas distribués indépendamment lors de la méiose.

La recombinaison produit des génotypes avec de nouvelles combinaisons des allèles parentaux.

Deux chromosomes homologues peuvent échanger des parties d'eux-mêmes par crossing-over.

La recombinaison résulte soit d'un assortiment indépendant (fréquence de recombinaison de 50 pour cent) soit de crossing-over (fréquence de recombinaison comprise entre 0 et 50 pour cent).

Les locus des gènes situés sur un même chromosome peuvent être cartographiés les uns par rapport aux autres, en mesurant les fréquences de recombinants produits par crossing-over.

Les distances génétiques basées sur les mesures de recombinaison sont approximativement additives.

Les crossing-over se produisent de façon exceptionnelle pendant la mitose.

Un crossing-over résulte de la formation d'un segment d'ADN hétérologue (hétéroduplex) — un brin provenant d'une chromatide parentale et l'autre brin provenant de l'autre chromatide.

Des chiasmas, les manifestations visibles de crossing-over au cours de la méiose dans un testicule de sauterelle.

(John Cabisco/Visuals Unlimited)

La variation héréditaire est un élément clé de la génétique. Deux processus sont responsables de la variation génétique, la *recombinaison* (le sujet de ce chapitre) et la *mutation*. La mutation est la source fondamentale de changement génétique ; de nouveaux allèles apparaissent dans tous les organismes, certains spontanément, d'autres à la suite d'une exposition à des agents mutagènes de l'environnement. Ces nouveaux allèles constituent le matériel de base pour un second niveau de variation, qui résulte de la recombinaison. Comme son nom le suggère, la recombinaison est le processus par lequel les allèles de gènes différents s'associent en de nouvelles combinaisons. Chez les Eucaryotes, la plupart des recombinaisons se produisent lors de la méiose. La recombinaison méiotique brasse les gènes en combinant différemment des paires d'allèles hétérozygotes dans les produits de la méiose (les ovules et les spermatozoïdes chez les animaux et les plantes). Si, pour la clarté de l'illustration, nous supposons qu'un organisme diploïde est hétérozygote pour seulement 10 gènes (une grosse sous-estimation par rapport à la plupart des organismes), alors le nombre total de génotypes possibles pour les cellules sexuelles est de $2^{10} = 1024$. Les processus cellulaires responsables de la recombinaison ont également lieu lorsque les allèles sont homozygotes, mais alors, l'événement de recombinaison ne peut être détecté au niveau génétique. On devrait toujours se rappeler que le généticien dépend de l'hétérozygotie — c'est-à-dire de différences génétiques — pour pouvoir effectuer une analyse génétique.

Dans la nature, mutation et recombinaison assurent la variation héréditaire qui est le matériel de base de la sélection naturelle. Cependant, les généticiens utilisent la mutation et la recombinaison comme matériel pour l'analyse génétique expérimentale. La recombinaison est très utile pour créer des génotypes spécifiques nécessaires à des fins expérimentales, en combinant les allèles présents dans les lignées génétiques établies. Parfois, ces allèles intéressants ont été isolés dans la nature, mais le plus souvent ils ont été induits expérimentalement par analyse de mutations. La recombinaison sert également à déterminer sur quel chromosome est situé un gène intéressant et à cartographier la position spécifique de celui-ci sur le chromosome. Ces deux utilisations sont des aspects essentiels de la génétique.

Comment peut-on définir la recombinaison ? Nous avons vu qu'il s'agissait d'un processus qui consiste à rassembler des allèles dans de nouvelles combinaisons. Toutefois, les généticiens la définissent d'une façon précise qui peut s'appliquer expérimentalement. Puisque la recombinaison chez les Eucaryotes est en grande partie un processus méiotique, pour le moment nous pouvons la définir en terme de méiose, comme l'indique le message suivant.

MESSAGE

La recombinaison caractérise tout processus (chez une cellule germinale ou méiocyte), engendrant un produit haploïde de méiose, dont le génotype est différent de ceux des deux génotypes haploïdes à l'origine du méiocyte diploïde. Le produit de la méiose ainsi formé est généralement appelé un *recombinant*.

Cette définition souligne le fait que l'on détecte la recombinaison en comparant les génotypes *résultants* de la méiose avec les génotypes *initiaux* (figure 5-1). Les génotypes initiaux sont les deux génotypes haploïdes qui se sont combinés pour former le méiocyte, la cellule diploïde qui va subir la méiose. Chez une plante ou un animal, on s'intéresse aux gamètes : les combinaisons alléliques des gamètes produits par un organisme (résultants) sont comparées aux combinaisons alléliques dans les gamètes qui ont fusionné pour créer l'organisme de départ (génotypes initiaux). Dans l'exemple de la figure 5-1, les génotypes initiaux sont *AB* et *ab* et les génotypes résultants sont *AB* et *ab* ainsi que deux nouveaux génotypes, *Ab* et *aB*, qui sont par définition des recombinants.

C'est dans les cycles vitaux des champignons haploïdes que la recombinaison est illustrée le plus simplement. Prenons comme exemple le croisement de cultures de deux mutants différents du champignon *Neurospora*. L'un est appelé givré, un phénotype mutant d'un allèle *g* qui entraîne un type de croissance hautement ramifiée qui ressemble à du givre sur une vitre. L'allèle de type sauvage *g*⁺ entraîne une croissance étalée moins ramifiée. L'autre mutant est jaune, à cause de l'allèle *j*. L'allèle de type sauvage *j*⁺ entraîne la production de pigment orange. Le croisement peut être désigné par :

g j⁺ (givrée, orange) x *g*⁺ *j* (étalée, jaune)

Les cellules de ces cultures s'unissent pour former un méiocyte diploïde — elles représentent les gènes initiaux de la méiose. La méiose se déroule et entraîne la formation de produits haploïdes de méiose appelés *ascospores*,

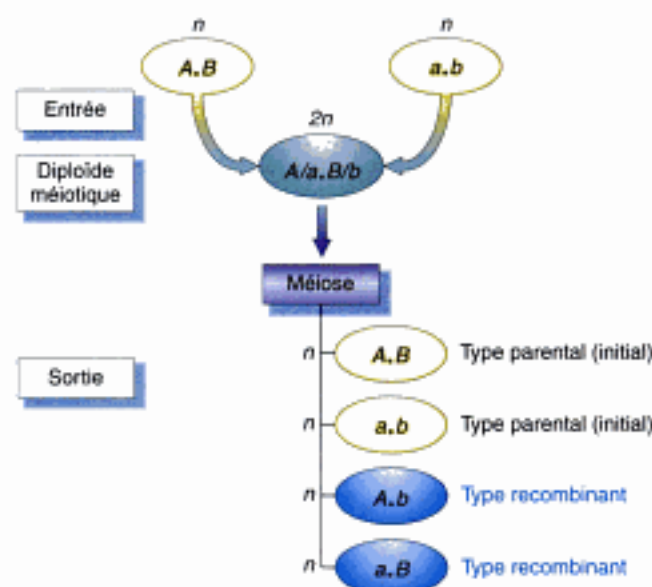


Figure 5-1 Les recombinants sont ces produits de la méiose dont les combinaisons alléliques diffèrent de celles des cellules haploïdes à l'origine du diploïde méiotique.

selon le cycle vital haploïde standard de ce champignon. Les phénotypes des cultures poussant à partir des ascospores seront :

Givrée, orange	ff^+
Étalée, jaune	f^+j
Étalée, orange	f^+j^+
Givrée, jaune	fj

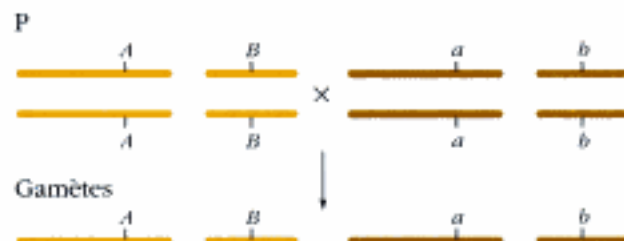
D'après la définition de la recombinaison, les descendants de phénotypes étalé, orange et givré, jaune sont des recombinants — ils représentent de nouvelles combinaisons d'allèles par rapport aux génotypes des parents.

Il existe deux processus méiotiques qui entraînent une recombinaison : il s'agit de l'**assortiment indépendant** des gènes sur des paires différentes de chromosomes et du **crossing-over** entre des gènes situés sur la même paire de chromosomes. Nous en parlerons dans les deux paragraphes suivants.

L'ASSORTIMENT INDÉPENDANT

Premièrement, puisque nous allons commencer à analyser deux gènes ou plus dans des croisements, il nous faut enrichir notre symbolisme. Un symbolisme utile, largement utilisé par les généticiens pour représenter deux gènes situés sur des chromosomes différents consiste à les écrire séparés par un point-virgule. Voici deux exemples provenant de diploïdes, $A/a ; B/b$ et $m^+/m ; p^+/p^+$. Voici deux exemples d'haploïdes : let^+ ; ade et fr^+ ; trp^+ .

Nous pouvons illustrer le processus d'assortiment indépendant par un exemple imaginaire, disons chez des souris, qui met en jeu les paires d'allèles A/a et B/b . Nous supposons que ces gènes sont situés sur des paires différentes de chromosomes homologues, une paire courte et une paire longue. Nous allons tout d'abord croiser deux lignées pures pour créer un **dihybride**, ou double hétérozygote, chez lequel nous étudierons la recombinaison. Les deux lignées que nous allons croiser sont $A/A ; B/B$ et $a/a ; b/b$:



La F_1 ne peut être que d'un seul type :



Nous pouvons à présent utiliser la loi de ségrégation égale pour prédire les génotypes des gamètes produits par la F_1 et leurs proportions. Nous savons que $1/2$ des gamètes seront A et $1/2$ a . Nous savons également que $1/2$ des gamètes seront B et $1/2$ b . Ces deux rapports doivent être multipliés l'un par l'autre pour déterminer l'ensemble des génotypes des gamètes. Pourquoi ? Rappelez-vous que d'après la règle du produit (chapitre 4), la probabilité que des événements indépendants se produisent simultanément est égale au produit de chacune de leurs probabilités. Donc, puisque les deux gènes en question sont indépendants, il faut multiplier les probabilités des gamètes. Les fréquences de $1/2$ attendues pour les gamètes sont effectivement des probabilités de $1/2$ (ou 50 pour cent) ; les gamètes seront donc répartis de la façon suivante :



La figure 5-2 à la page suivante montre la façon dont ce rapport est produit au niveau des déplacements des chromosomes lors de la méiose. Cela s'explique de façon simplifiée par le fait qu'il y a deux façons différentes mais aussi fréquentes l'une que l'autre, que les fibres du fuseau se fixent aux centromères :

- Dans la moitié des méiocytes, A et B sont tirés vers un pôle et a et b vers l'autre pôle.
- Dans l'autre moitié des méiocytes, A et b sont tirés vers un pôle et a et B vers l'autre pôle.

Le principe général est connu sous le nom d'**assortiment indépendant** des paires d'allèles. Ce principe est basé sur le fait que la ségrégation égale d'une paire d'allèles est *indépendante* de la ségrégation égale de l'autre paire d'allèles car ils sont situés sur des chromosomes différents.

MESSAGE

L'assortiment des paires d'allèles situés sur des paires de différents chromosomes est indépendant.

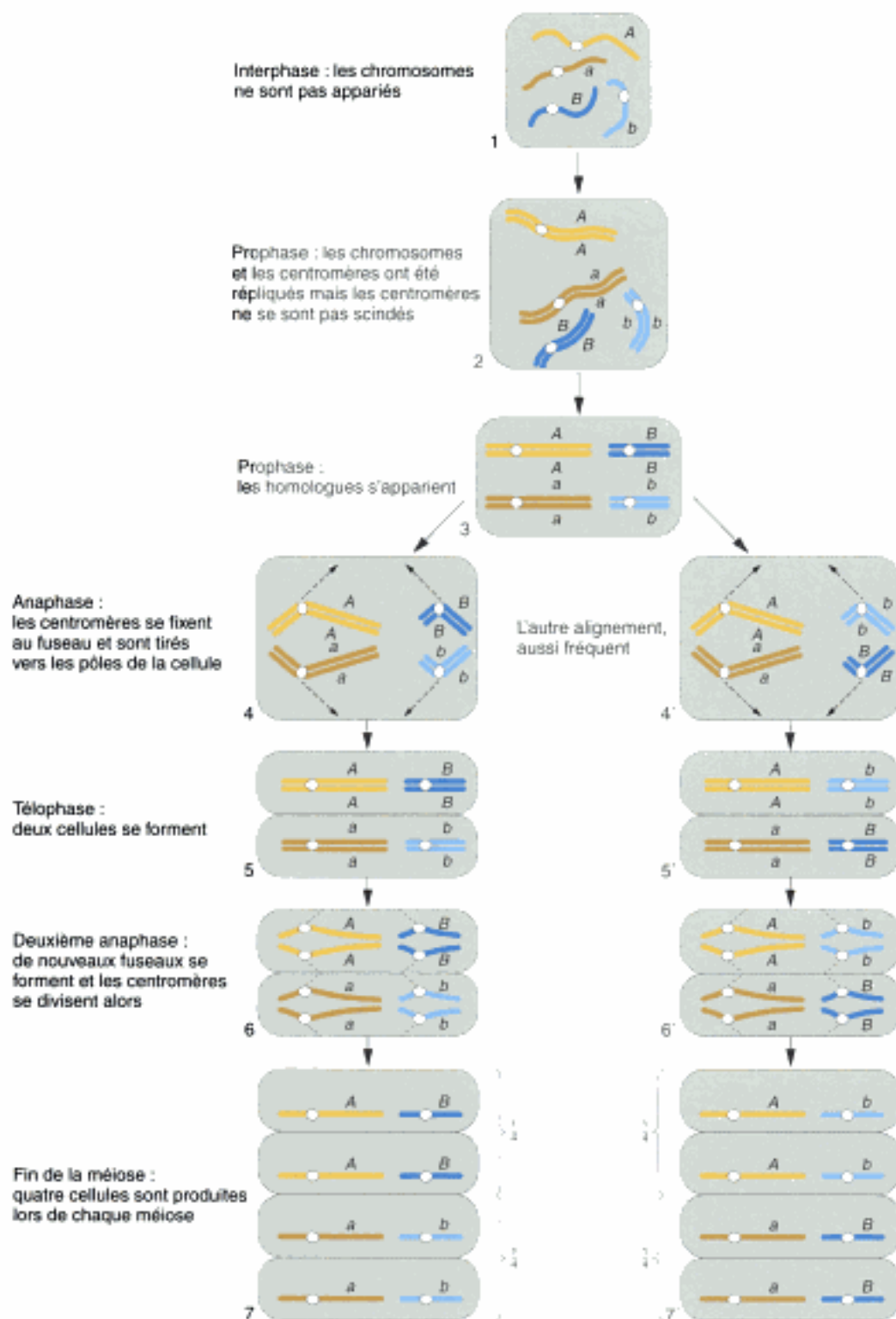


Figure 5-2 Une méiose dans une cellule diploïde de génotype $A/a; B/b$, montrant de quelle façon la ségrégation et l'assortiment de paires différentes de chromosomes donnent lieu à un rapport gamétique 1 : 1 : 1 : 1.

Hidden page

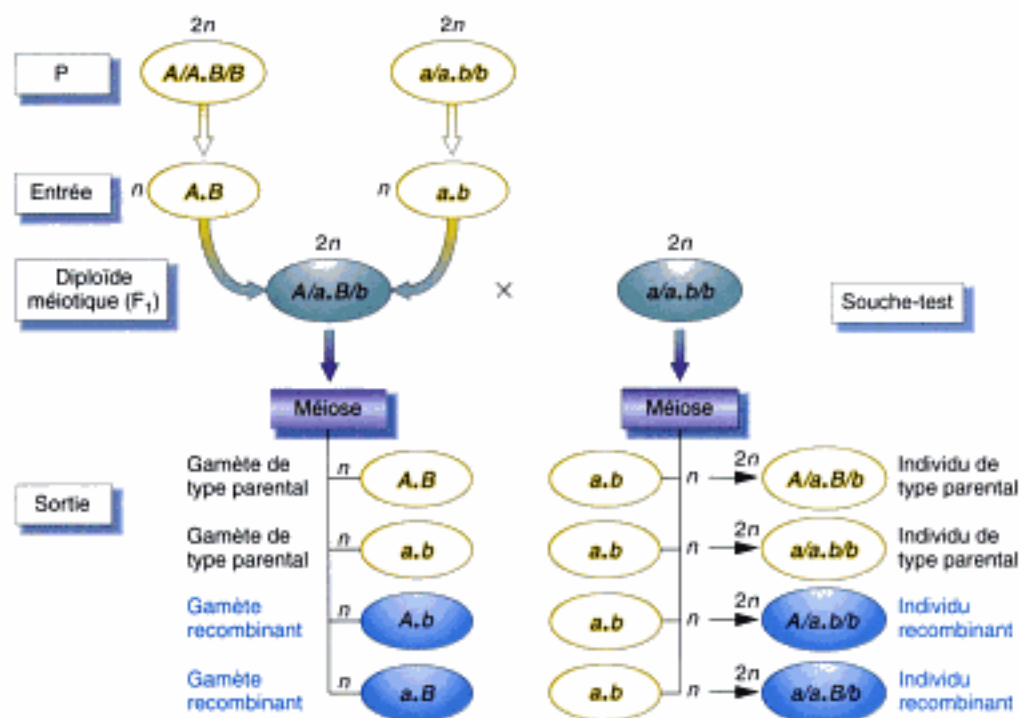
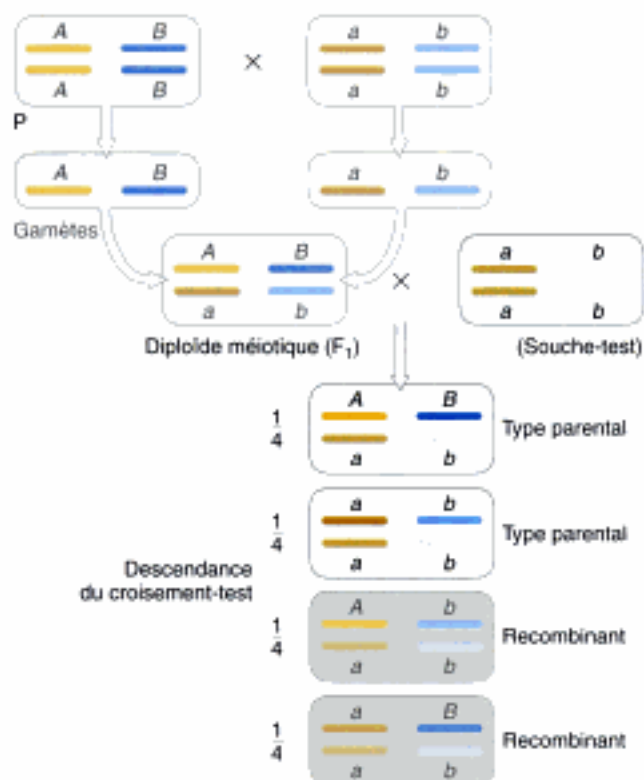


Figure 5-3 La détection de la recombinaison chez des organismes diploïdes. La manière la plus simple de détecter les produits recombinants d'une méiose chez un diploïde est de réaliser le croisement d'un hétérozygote avec une souche-test récessive.

Dans ces exemples, nous avons illustré l'assortiment indépendant avec des gènes connus pour être situés sur des



chromosomes différents. Les rapports standard tels que les rapports des descendants d'un croisement-test 1 : 1 : 1 : 1 peuvent aussi être utilisés pour *déduire* que l'assortiment des gènes est indépendant. Nous pouvons illustrer cela par un croisement utilisé par Mendel. Nous avons vu au chapitre 4 que Mendel avait déduit que les deux phénotypes de couleurs de pois jaune et vert étaient déterminés par deux allèles d'un même gène, J (jaune) et j (vert). Par d'autres analyses, il apprit que les phénotypes de l'aspect des pois, lisse ou ridé, étaient déterminés par les allèles d'un autre gène, avec R déterminant l'aspect lisse et r l'aspect ridé. Puisque dans ce cas nous ne savons pas si les deux gènes se trouvent sur des chromosomes séparés ou sur le même chromosome, nous devons trouver une façon de représenter symboliquement cette incertitude. Il n'y a pas de convention universelle pour cette situation, mais dans ce livre, nous utiliserons un *point* pour signifier cette incertitude. D'après son croisement entre une plante de lignée pure à pois lisses et jaunes ($R/R.J/J$) et une plante de lignée pure à pois ridés et verts ($r/r.j/j$), Mendel obtint une F_1 dihybride qui était (comme prévu) lisse,

Figure 5-4 L'assortiment indépendant qui produit toujours une fréquence de recombinaison de 50 pour cent. Ce schéma montre deux paires de chromosomes chez un organisme diploïde avec A et a sur une paire et B et b sur l'autre. Remarquez que nous aurions pu représenter la situation haploïde en enlevant la partie marquée P et le croisement-test.

Hidden page

Ce rapport de 9 : 3 : 3 : 1 est caractéristique de dihybrides autofécondés présentant un assortiment indépendant. En général, le rapport de la descendance de l'autofécondation d'un dihybride pour des gènes situés sur des chromosomes distincts peut être écrit ainsi :

Phénotype		
	Premier gène	Deuxième gène
$\frac{9}{16}$	Dominant	Dominant
$\frac{3}{16}$	Dominant	Récessif
$\frac{3}{16}$	Récessif	Dominant
$\frac{1}{16}$	Récessif	Récessif

Une photographie d'un rapport phénotypique 9 : 3 : 3 : 1 chez le maïs (*Zea mays*) est présentée dans la figure 5-6.

MESSAGE

Un rapport 1 : 1 : 1 : 1 obtenu lors d'un croisement-test réalisé avec un dihybride et un rapport 9 : 3 : 3 : 1 obtenu après autofécondation d'un dihybride reflètent tous deux un rapport de 1 : 1 : 1 : 1 chez les gamètes, qui montre que les paires d'allèles subissent un assortiment indépendant (généralement parce qu'elles sont situées sur des paires différentes de chromosomes) et que la FR est de 50 pour cent.



Figure 5-6 Un rapport 9 : 3 : 3 : 1 dans les phénotypes de grains de maïs. Chaque grain représente un descendant. La descendance résulte de l'autofécondation d'un individu de génotype $A/a ; B/b$, où A = violet, a = jaune, B = lisse et b = ridé. (Anthony Griffiths).

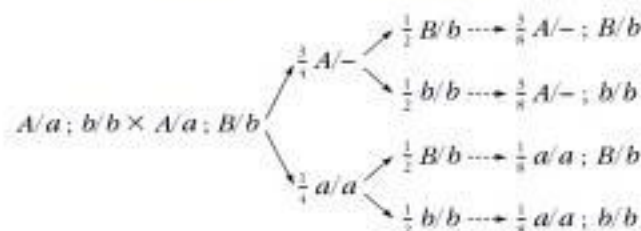
Le calcul des rapports phénotypiques et génotypiques pour des gènes dont l'assortiment est indépendant

Au chapitre 4, nous avons calculé les rapports phénotypiques attendus pour la descendance de croisements impliquant des allèles d'un même gène. Par exemple

$$A/a \times A/a \longrightarrow \frac{3}{4} A/- \text{ et } \frac{1}{4} a/a$$

$$A/a \times a/a \longrightarrow \frac{1}{2} A/a \text{ et } \frac{1}{2} a/a$$

Cette connaissance permet de prédire les génotypes et les phénotypes impliquant plusieurs gènes si leur assortiment est indépendant. Par exemple, dans le croisement $A/a ; b/b \times A/a ; B/b$, nous pourrions vouloir calculer le rapport génotypique attendu pour la descendance. Nous savons que dans la descendance, le rapport pour le premier gène sera $3/4 A/- ; 1/4 a/a$ et pour le deuxième $1/2 B/b ; 1/2 b/b$. Donc, si ces gènes subissent un assortiment indépendant, nous pouvons combiner les deux rapports phénotypiques au hasard en faisant un schéma que l'on appelle un **arbre**. Dans l'arbre, les rapports des descendants sont obtenus simplement en multipliant les proportions d'après la loi du produit, comme on le voit ci-dessous :



On pourrait obtenir les mêmes rapports chez les descendants en utilisant un tableau, qui présente également les associations aléatoires. Remarquez que les axes de ce tableau indiquent les proportions phénotypiques. (Les tableaux sont très utiles en génétique et peuvent être utilisés de nombreuses façons).

		Descendants	
		Gène A/a	
		$\frac{3}{4} A/-$	$\frac{1}{4} a/a$
Descendants	Gène B/b	$\frac{3}{8} A/- ; B/b$	$\frac{3}{8} B/b ; a/a$
	$\frac{1}{2} b/b$	$\frac{3}{8} A/- ; b/b$	$\frac{1}{8} a/a ; b/b$

Dans un croisement plus complexe, disons, $A/a ; B/b ; C/c \times a/a ; B/b ; C/c$, nous aimerions pouvoir prédire la proportion d'un certain génotype spécifique de la descendance, nécessaire pour une expérience, par exemple, le génotype $a/a ; b/b ; c/c$ pour l'utiliser comme souche-test. Puisque nous savons que les proportions génotypiques pour chacun des gènes seront les suivants :

$$\begin{aligned} A/a \times a/a &\longrightarrow \frac{1}{2} A/a + \frac{1}{2} a/a \\ B/b \times B/b &\longrightarrow \frac{1}{4} B/B + \frac{1}{2} B/b + \frac{1}{4} b/b \\ C/c \times C/c &\longrightarrow \frac{1}{4} C/C + \frac{1}{2} C/c + \frac{1}{4} c/c \end{aligned}$$

Si l'assortiment des gènes est indépendant, nous pouvons utiliser simplement la loi du produit pour multiplier les proportions de chacun des génotypes *qui nous intéressent* (indiqués en gras) pour obtenir la proportion attendue (probabilité) de $a/a ; b/b ; c/c$; ce sera $1/2 \times 1/4 \times 1/4 = 1/32$. Donc, si nous avons besoin de ce génotype, il nous faudra examiner plus de 32 descendants pour avoir une chance raisonnable d'en obtenir un qui soit $a/a ; b/b ; c/c$.

Ces calculs simples, basés sur l'indépendance, peuvent être utilisés pour prédire les proportions des phénotypes, des génotypes ou des gamètes dans des croisements *si l'on sait* ou *si l'on suppose* que l'assortiment des gènes est indépendant. Cependant, comme nous l'avons vu au chapitre 4 et plus haut dans ce chapitre, l'analyse génétique fonctionne dans les deux sens. Nous pouvons donc également utiliser les proportions spécifiques des descendants pour arriver à des déductions sur les génotypes et les phénotypes des parents, si on ne les connaît pas. Par exemple, que pourrions-nous déduire si nous observons les nombres suivants de phénotypes dans la descendance d'une autofécondation de plante de phénotype normal sauvage à grands pétales bleus :

Bleu, grand	182
Bleu, petit	60
Blanc, grand	57
Blanc, petit	21

Nous remarquerions que les nombres de descendants sont très proches du rapport 9 : 3 : 3 : 1. Nous pourrions donc déduire que :

1. Le parent devait être hétérozygote pour deux gènes affectant la couleur et la taille des pétales.
2. La couleur bleue des pétales est dominante sur le blanc et la grande taille des pétales est dominante sur la petite taille. Nous pourrions inventer les symboles d'allèles suivants : b^* = bleu et b = blanc, et p^* = grand et p = petit.
3. L'assortiment des deux gènes est indépendant et ils sont fort probablement situés sur des paires différentes de chromosomes ; la plante autofécondée devait donc être de génotype $b^*/b ; p^*/p$.

LE CROSSING-OVER

Le deuxième grand moyen de produire des recombinants est le crossing-over. Lorsque des paires homologues de chromatides sœurs sont observées dans leur configuration « empilée » lors de la méiose, on observe des structures en forme de croix appelées **chiasm**, formées entre des chromatides non sœurs de paires d'homologues. Chaque paire d'homologues présente un ou plusieurs chiasm. Les chiasm peuvent se former à n'importe quel endroit d'une paire mais leur position est variable selon les méiocytes. Les chiasm semblent se produire plus ou moins au hasard. Ces chiasm correspondent aux positions des **crossing-over**, des endroits où un événement de cassure et de réunion a eu lieu entre des chromatides non sœurs homologues (voir Les grandes découvertes de la génétique 5-1 à la page suivante). Ce processus ressemble un peu au montage d'un film ; une coupure est réalisée à une position identique sur deux homologues, puis les homologues opposés sont épissés ensemble (raboutés). La figure 5-7 montre deux chiasm et les crossing-over qui leur sont associés.

Les gènes situés sur le même chromosome sont dits **liés** pour la raison évidente qu'ils sont liés physiquement ou reliés l'un à l'autre par le segment de chromosome situé entre eux. Un gène se trouve toujours à la même place (appelée **locus**) sur un chromosome. N'importe quel allèle du gène sera donc toujours au même locus. Par conséquent, en décrivant deux gènes situés sur le même chromosome, nous pourrions dire que :

- Les deux gènes sont liés.
- Les allèles des deux gènes sont liés.
- Les locus des deux gènes sont liés.

— tout ceci signifie la même chose.

Lorsque les gènes sont liés chez un dihybride, un crossing-over peut conduire à la production de recombinants. Pour illustrer ce processus, il nous faut introduire de nouveaux termes. Il existe deux sortes de dihybrides mettant en jeu des gènes liés. La différence réside dans le fait que les allèles mutants des deux gènes sont situés ou non sur le même homologue parental (dihybride *cis*) ou sur des homologues parentaux différents (dihybride *trans*). En supposant que les allèles récessifs sont les

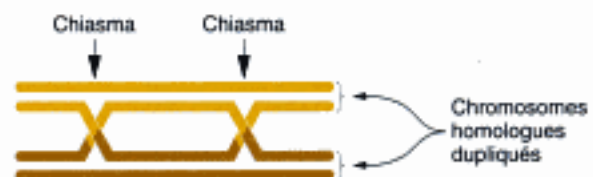


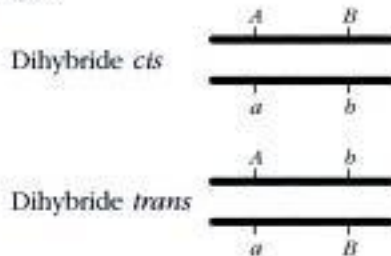
Figure 5-7 Représentation schématique des chiasm lors de la méiose. Chaque ligne représente une chromatide appartenant à une paire de chromosomes appariés.

Hidden page

Figure 5-8 Chez des dihybrides pour des gènes liés, les recombinants apparaissent à la suite de méioses au cours desquelles des chromatides non sœurs subissent des *crossing-over* entre les gènes étudiés.

	Chromosomes méiotiques	Produits de la méiose	
Méioses sans <i>crossing-over</i> entre les gènes			Parental
			Parental
			Parental
			Parental
Méioses avec un <i>crossing-over</i> entre les gènes			Parental
			Recombinant
			Recombinant
			Parental

allèles mutants, les deux arrangements se présentent de la façon suivante :



Les gènes liés sont symbolisés ainsi : les allèles présents sur un homologue sont écrits à gauche de la barre oblique (/) et ceux situés sur l'autre homologue, à droite de cette barre. Les gènes sont indiqués dans le même ordre des deux côtés de la barre oblique. Les dihybrides *cis* et *trans* ci-dessus pourraient être représentés par *AB/ab* et *Ab/aB*, respectivement. (Remarquez qu'il n'y a pas de signe de ponctuation entre des allèles situés sur le même homologue).

Comment pouvons-nous dire sur lequel de ces deux arrangements nous travaillons (*cis* ou *trans*) — ils contiennent les mêmes allèles et sont parfaitement identiques du point de vue de leur phénotype ? Il existe deux façons de les distinguer. Premièrement, les deux parents du dihybride sont souvent des souches de lignée pure et ceci dicte automatiquement l'arrangement allélique de leur descendance. Par exemple, si les deux parents d'un dihybride sont *Ab/Ab* et *aB/aB*, leurs gamètes doivent être *Ab* et *aB* et le dihybride doit donc être *Ab/aB* (*trans*). À l'inverse, si les parents sont *AB/AB* et *ab/ab*, les descendants dihybrides doivent être en *cis*, *AB/ab*. Une autre façon de déterminer si un dihybride pour des gènes liés est en *cis* ou en *trans* est de travailler à partir des rapports de la descendance, comme nous allons le voir plus loin.

En utilisant un dihybride *cis*, regardons de quelle façon un *crossing-over* conduit à des recombinants. Les *crossing-over* se produisent plus ou moins au hasard le long des homologues appariés, dans la population des méiocytes en cours de méiose, un *crossing-over* se

produira entre les deux gènes chez certains et d'autres homologues n'en subiront pas. L'ensemble des quatre produits de la méiose sera différent dans chaque cas, comme le montre la figure 5-8. Dans les méioses avec un *crossing-over*, deux des produits haploïdes de méiose seront recombinants, de génotypes *Ab* et *aB*. (S'il s'agissait d'un dihybride *trans*, les recombinants seraient *AB* et *ab*). La figure 5-9 représente la descendance typique d'un dihybride *cis*, qu'il faudrait comparer avec la figure 5-4, représentant la situation comparable pour un assortiment indépendant.

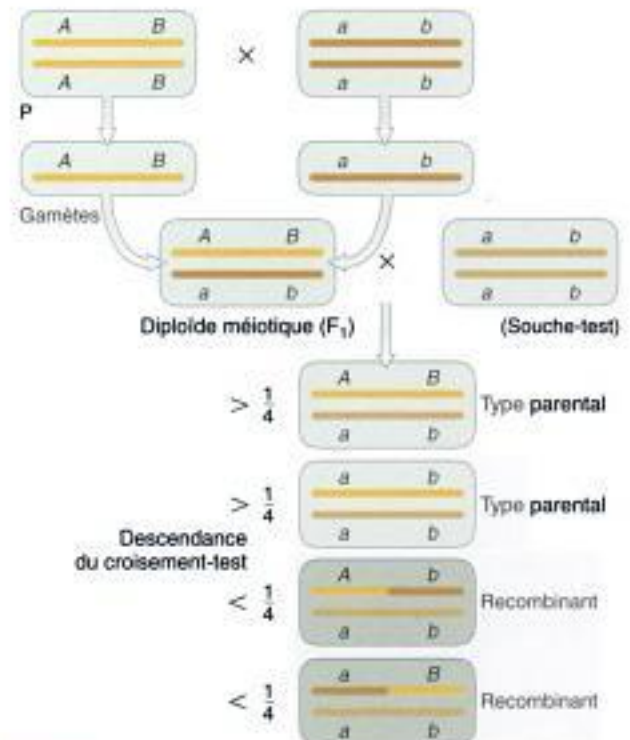


Figure 5-9 Recombinants produits par *crossing-over*. Remarquez que les fréquences de recombinaison s'ajoutent, mais que leur somme reste inférieure à 50 pour cent.

Hidden page

nombre d'analyses par recombinaison augmente, mettant en jeu de plus en plus de gènes, le chromosome entier « s'étioffe ». C'est seulement après avoir placé sur la carte, les gènes situés à proximité de chaque extrémité du chromosome et après avoir déterminé les positions des gènes par analyse des réarrangements chromosomiques (chapitre 8), que les positions absolues des gènes sur les chromosomes peuvent être attribuées.

Comme nous l'avons dit plus tôt, l'analyse génétique fonctionne en deux sens. Dans le cas de l'analyse des liaisons génétiques, nous avons vu qu'en utilisant les mesures de FR, nous pouvions dessiner une carte génétique. À l'inverse, si nous disposons de la carte, nous pouvons prédire les fréquences des descendants dans les différentes catégories de génotypes. Par exemple, si nous savons que chez la drosophile, les gènes autosomiques pour la couleur de l'œil (P = rouge, p = pourpre) et de la taille des ailes (V = longues, v = vestigiales) sont liés sur le même chromosome et distants de 11 u.g., et si nous réalisons un croisement-test entre un dihybride *trans* Pv/pV et pv/pv , nous pouvons prédire qu'il y aura 11 pour cent de recombinants. De plus, il devrait y avoir 5,5 pour cent de PV/pv et de pV/pv car les deux types de recombinants sont des produits *réciroques* du même type d'événement de crossing-over. Les 89 pour cent restants ne sont pas recombinants et correspondent pour 44,5 pour cent à chacun des génotypes Pv/pv et pV/pv .

MESSAGE

La recombinaison entre des gènes liés peut être utilisée pour cartographier leur éloignement sur le chromosome.

L'unité génétique (1 u.g.) est définie comme une fréquence de recombinaison de 1 pour cent.

Chez un dihybride pour des locus liés, la fréquence de recombinaison ne dépasse jamais 50 pour cent, quel que soit l'éloignement des deux locus. Cela n'est pas intuitif car on imagine que plus la distance chromosomique augmente, plus il y a de crossing-over et donc plus il y aura de recombinants. Cependant, on peut montrer simplement que dans les méiocytes comportant *n'importe quel* nombre de crossing-over autre que zéro, la proportion moyenne de recombinants est de 50 pour cent. Donc le facteur déterminant de la fréquence de recombinaison est la proportion relative de la classe de zéro crossing-over. Pour les régions chromosomiques longues, il n'y aura pas de classe zéro et la FR sera de 50 pour cent mais ne pourra jamais dépasser ce nombre.

MESSAGE

Chez un dihybride pour des gènes liés, la FR sera comprise entre 0 et 50 pour cent.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 5-2

La première carte chromosomique

Alfred Sturtevant commença sa carrière en génétique en 1911 comme étudiant de deuxième cycle dans le laboratoire du grand généticien Thomas Hunt Morgan. Morgan confia à Sturtevant la tâche d'expliquer certaines données. Celles-ci concernaient différentes combinaisons alléliques qui apparaissaient dans des croisements de gènes présentant une transmission liée au chromosome X. En une soirée, Sturtevant exploita toutes les données et ce faisant, développa la méthode analytique permettant de cartographier les gènes sur les chromosomes, qui est toujours en vigueur aujourd'hui. En se remémorant l'événement plus tard, il dit :

À la fin de l'année 1911, en parlant avec Morgan, je réalisai soudain que les variations dans l'intensité de la liaison génétique (association entre les allèles de différents gènes) déjà attribuées par Morgan à la séparation spatiale des gènes, offraient la possibilité de déterminer des séquences linéaires des gènes sur un chromosome. Je rentrai chez moi et passai la plus grande partie de la nuit (négligeant mon travail d'étudiant) à élaborer la première carte chromosomique.

L'analyse fut publiée en 1913. L'introduction disait :

Le parallèle entre le comportement des chromosomes en réduction, et celui des facteurs de Mendel lors de la ségrégation a été souligné pour la première fois par Sutton (1902), bien que plus tôt dans la même année Boveri (1902) ait fait référence à une association possible.

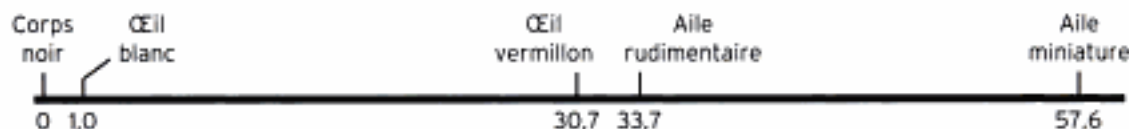
L'article se poursuivait par la cartographie de six gènes liés à l'X, faisant appel à la logique suivante :

La proportion de crossing-over pourrait être utilisée comme un indicateur de la distance séparant n'importe quelle paire de facteurs. Puis, en déterminant les distances (dans le sens ci-dessus) entre A et B et entre B et C, on devrait pouvoir prédire AC. Car, si les proportions de crossing-over représentent réellement la distance, AC doit être approximativement égale à AB plus BC ou à AB moins BC, et non pas à une valeur intermédiaire.



Robert Galentine/Visuals Unlimited

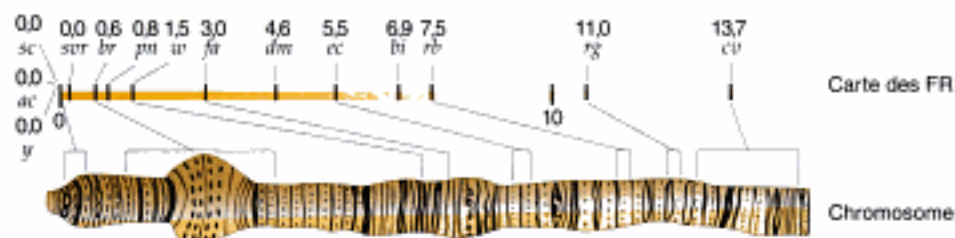
La carte du chromosome X de Sturtevant (la première carte génétique) était la suivante. Le zéro correspondait arbitrairement à une extrémité déterminée de la carte :



Remarquez que (par exemple) la distance génétique entre le gène de l'aile rudimentaire et celui de l'aile miniature est de 57,6 moins 33,7, soit 23,9 unités génétiques.

La conclusion de Sturtevant était la suivante :

Les résultats apportent un nouvel argument en faveur de la conception chromosomique de l'hérédité, car ils indiquent fortement que les facteurs étudiés sont arrangés de façon linéaire, au moins du point de vue mathématique.



Réaliser une carte à l'aide du croisement-test d'un trihybride

Jusqu'ici, nous avons considéré l'analyse de la recombinaison chez des dihybrides. Nous avons vu qu'en rassemblant les distances génétiques obtenues par de nombreux croisements-test effectués entre des dihybrides, il était possible de dresser des cartes complètes de liaison génétique qui correspondent aux cartes des locus sur les chromosomes. Cependant, dans certaines situations, les chercheurs disposent d'hétérozygotes multiples (trihybrides, tétrahybrides, etc.) et dans ces cas, il est possible de réaliser plusieurs calculs de distance génétique lors d'un seul croisement-test avec un récessif multiple. Nous allons suivre un exemple en utilisant le croisement-test d'un trihybride, parfois appelé **croisement-test à trois points**.

L'exemple analyse les allèles au niveau de trois locus chez la drosophile :

1. *v* (yeux vermillons) contre *v⁺* (yeux rouges)
2. *cv* (sans veine transversale sur l'aile) contre *cv⁺* (présence d'une veine transversale)

3. *ct* (bout des ailes incisé) contre *ct⁺* (bout des ailes lisse)

Au départ, nous ignorons si ces gènes sont liés, c'est pourquoi ils ne doivent pas être écrits dans un ordre particulier. Les souches parentales disponibles sont des mouches homozygotes double récessives, de génotype *v/v cv/cv ct/ct* et des mouches homozygotes simple récessives de génotype *v/v ct⁺/ct⁺ ct⁺/ct⁺*. À partir de ce croisement, on obtient une descendance trihybride de génotype *v/v⁺ cv/cv⁺ ct/ct⁺* et les femelles de ce génotype subissent un croisement-test avec des triple récessifs de génotype *v/v cv/cv ct/ct*. Comme il y a trois paires d'allèles, il y a $2 \times 2 \times 2 = 8$ génotypes gamétiques possibles obtenus à partir de la femelle. Comme pour les croisements-test plus simples, ces génotypes gamétiques déterminent les huit types de descendants issus de ce croisement-test. Par souci de simplicité, les descendants sont écrits suivant leur génotype gamétique et la contribution *v cv ct* de la souche-test n'est pas indiquée. Les données sont présentées ci-dessous sous la forme des nombres de descendants dans un échantillon total de 1448 mouches obtenues après un croisement réel :

<i>v</i>	<i>cv</i> ⁺	<i>ct</i> ⁺	580
<i>v</i> ⁺	<i>cv</i>	<i>ct</i>	592
<i>v</i>	<i>cv</i>	<i>ct</i> ⁺	45
<i>v</i> ⁺	<i>cv</i> ⁺	<i>ct</i>	40
<i>v</i>	<i>cv</i>	<i>ct</i>	89
<i>v</i> ⁺	<i>cv</i> ⁺	<i>ct</i> ⁺	94
<i>v</i>	<i>cv</i> ⁺	<i>ct</i>	3
<i>v</i> ⁺	<i>cv</i>	<i>ct</i> ⁺	5
			1448

Il nous faut utiliser l'approche standard de la recombinaison, mais nous devons faire attention à notre classification des types parentaux et recombinants. Remarquez que les génotypes parentaux pour les triple hétérozygotes sont *v*⁺ *cv* *ct* et *v* *cv*⁺ *ct*⁺. Nous devons prendre cela en considération lorsque nous décidons ce qui constitue un recombinant. Les données sont analysées au niveau de deux locus à la fois.

En commençant par les locus *v* et *cv*, nous savons que les recombinants doivent être de génotype *v* *cv* et *v*⁺ *cv*⁺ et qu'ils sont 45 + 40 + 89 + 94 = 268. Cela représente une FR de 18,5 pour cent parmi les 1448 mouches.

Pour les locus *v* et *ct*, les recombinants sont *v* *ct* et *v*⁺ *ct*⁺. Ils sont 89 + 94 + 3 + 5 = 191, soit une FR de 13,2 pour cent.

Pour *ct* et *cv*, les recombinants sont *cv* *ct*⁺ et *cv*⁺ *ct*. Ils sont 45 + 40 + 3 + 5 = 93 parmi les 1448, soit une FR de 6,4 pour cent.

À l'évidence, tous les locus sont liés sur le même chromosome, car les valeurs de FR sont toutes nettement inférieures à 50 pour cent. Les locus *v* et *cv* présentant la valeur de FR la plus importante, ils doivent être les plus éloignés. Le locus *ct* doit donc être situé entre eux. On peut dessiner une carte de la façon suivante



Il faut noter plusieurs points importants ici. Le premier est que nous avons déduit de ces résultats un ordre des gènes différents de celui utilisé lorsque nous avons dressé la liste des génotypes des parents et des descendants. Puisque le but de l'exercice était de déterminer les relations de liaison génétique entre ces gènes, l'énoncé original était nécessairement arbitraire ; on ne connaissait tout simplement pas l'ordre des gènes avant d'avoir analysé les données.

Deuxièmement, nous avons établi sans doute possible que *ct* se trouvait entre *v* et *cv* et nous avons déterminé les distances qui séparent *ct* des autres locus en unités génétiques. Mais nous avons arbitrairement placé *v* à gauche et *cv* à droite. La carte pourrait aussi bien être inversée.

Un troisième point à prendre en compte est que la somme des deux distances génétiques les plus faibles, 13,2 u.g. et 6,4 u.g. est de 19,6 u.g., ce qui est supérieur à 18,5 u.g., la distance calculée pour *v* et *cv*. Comment cela s'explique-t-il ? La réponse à cette question réside dans la

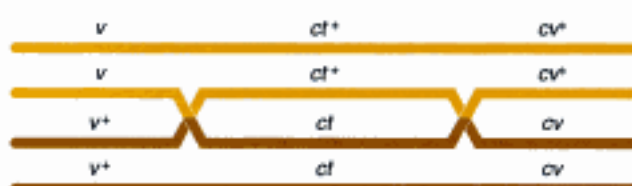


Figure 5-11 Exemple de double crossing-over. Remarquez qu'un double crossing-over produit des chromatides double recombinants qui maintiennent les combinaisons alléliques parentales au niveau des locus extérieurs.

façon dont nous avons analysé les deux classes les plus rares dans notre classification des recombinants pour les locus *v* et *cv*. Maintenant que nous avons la carte, nous pouvons voir que ces deux classes rares sont en fait des double recombinants, qui apparaissent à la suite de deux crossing-over (figure 5-11). Pourtant, nous n'avons pas compté les génotypes *v* *ct* *cv*⁺ et *v*⁺ *ct*⁺ *cv* lorsque nous avons calculé la valeur de la FR pour *v* et *cv* ; après tout, par rapport à *v* et à *cv*, ils correspondent à des combinaisons parentales (*v* *cv*⁺ et *v*⁺ *cv*). À la lumière de notre carte, nous voyons cependant que ceci a conduit à une sous-estimation de la distance séparant les locus *v* et *cv*. Non seulement nous aurions dû compter les deux classes les plus rares, mais nous aurions dû compter chacune d'elles *deux fois* car elles représentent deux classes de *double* recombinants. Nous pouvons donc corriger la valeur en ajoutant les nombres 45 + 40 + 89 + 94 + 3 + 3 + 5 + 5 = 284. Sur un total de 1448, cela représente exactement 19,6 pour cent, ce qui est identique à la somme des valeurs des deux composants.

Maintenant que nous avons exploité quelque peu les données de ce croisement, nous pouvons regarder en arrière dans la liste des descendants et voir qu'il est généralement possible de déduire l'ordre des gènes par leur examen, sans effectuer d'analyse de fréquence de recombinaison. Seuls trois ordres de gènes sont possibles, chacun avec un gène différent au milieu. Il est généralement vrai que les classes de double recombinants sont les plus petites. Seul un ordre devrait être compatible avec les classes les plus petites ayant été formées par deux crossing-over, comme le montre la figure 5-12 à la page suivante. Seul un ordre donne des double recombinants de génotype *v* *ct* *cv*⁺ et *v*⁺ *ct*⁺ *cv*. Remarquez en passant que la possibilité de détecter des double crossing-over dépend du fait de disposer d'un gène hétérozygote entre les deux crossing-over. Si les mères de ces descendants n'avaient pas été hétérozygotes (*ct*/*ct*⁺) nous n'aurions jamais pu identifier les classes de double recombinants.

Enfin, remarquez une nouvelle fois que les cartes de liaison cartographient simplement les locus les uns par rapport aux autres, en utilisant des unités génétiques standard. Nous ne savons pas où sont situés les locus sur un chromosome — ni même sur quel chromosome ils se trouvent. La carte de liaison génétique est pour l'essentiel une construction abstraite qui peut être corrélée à un chromosome spécifique et à des régions spécifiques du chromosome, seulement en appliquant des types particuliers d'analyses cytogénétiques, comme nous le verrons au chapitre 8.

Ordres possibles des gènes	Chromatides double recombinants

Figure 5-12 Avec trois gènes, seuls trois ordres de gènes sont possibles. En considérant chacun d'entre eux, les double crossing-over créent des génotypes uniques pour les double recombinants. Seule la première possibilité est compatible avec les données du texte.

MESSAGE

Les croisements-test mettant en jeu des hétérozygotes multiples permettent d'évaluer les liaisons génétiques entre plusieurs locus, à partir d'un seul croisement.

L'interférence

La détection de classes de double recombinants montre qu'il doit y avoir des double crossing-over. Sachant cela, l'une des questions qui se pose est de savoir si les crossing-over qui ont lieu dans des régions chromosomiques adjacentes sont indépendants ou s'ils interagissent l'un avec l'autre d'une façon ou d'une autre. Par exemple, nous pourrions nous demander si un crossing-over dans une région affecte la probabilité qu'il y ait un crossing-over dans une région adjacente. Apparemment la réponse est souvent « oui » ; cette interaction s'appelle **l'interférence**.

On peut aborder cette analyse de la façon suivante. Si les deux crossing-over dans les deux régions sont indépendants, alors, d'après la règle du produit, la fréquence des double recombinants devrait être égale au produit des fréquences de recombinaison dans les régions adjacentes. Dans les données de recombinaison v - ct - cv , la valeur de FR de v - ct est de 0,132 et la valeur de ct - cv est de 0,064. Donc s'il y a indépendance, les double recombinants devraient être attendus à une fréquence de $0,132 \times 0,064 = 0,0084$ (0,84 pour cent). Dans l'échantillon des 1448 mouches, on attendrait un total de $0,0084 \times 1448 = 12$ double recombinants. Mais les données montrent que

seuls 8 double recombinants ont en réalité été observés. Si ce déficit de recombinants était observé régulièrement, cela nous montrerait que les deux régions ne sont pas indépendantes et suggérerait que la distribution des crossing-over favorise les crossing-over uniques aux dépens des doubles. En d'autres termes, il y a une sorte d'interférence : un crossing-over diminue la probabilité d'un crossing-over dans une région adjacente.

L'interférence est quantifiée en calculant d'abord un terme appelé le **coefficient de coïncidence (c.d.c.)**. Ce coefficient se calcule en faisant le rapport des double recombinants observés sur ceux attendus et en soustrayant cette valeur de 1. Donc

$$\begin{aligned} \text{Interférence} &= 1 - \text{c.d.c.} \\ &= 1 - \frac{\text{Double recombinants observés}}{\text{Double recombinants attendus}} \end{aligned}$$

Dans notre exemple $1 - \left(\frac{8}{12}\right) = \frac{4}{12} = \frac{1}{3}$, soit 33 pour cent.

Dans certaines régions, on n'observe jamais de double recombinants. Dans ces cas, c.d.c. = 0, donc $1 - 0 = 1$ et l'interférence est totale. La plupart du temps, les valeurs d'interférence que l'on rencontre en cartographiant les locus d'un chromosome sont comprises entre 0 et 1, mais dans certaines situations particulières, les double crossing-over observés excèdent les double crossing-over attendus, aboutissant à des valeurs d'interférence négatives.

On rencontre couramment le phénomène de l'interférence au cours de l'analyse de liaisons génétiques, mais on ne connaît aucun mécanisme moléculaire l'expliquant.

Quelques exemples de cartes de liaison génétique

Les cartes de liaison génétique sont le prélude à toute étude sérieuse de génétique expérimentale chez n'importe quel organisme. Les chromosomes de nombreux organismes ont été intensivement cartographiés de cette façon. Les cartes résultantes représentent une quantité importante d'analyse génétique, généralement obtenue par les efforts conjoints de groupes de recherche du monde entier.

Le niveau d'analyse que nous avons atteint dans notre discussion est bien illustré par les cartes de liaison génétique de la larve de la lucilie bouchère (*Cochliomyia vicina*). Le stade larvaire de cet insecte — le « ver » — parasite parfois les blessures de mammifères et est un animal particulièrement nuisible pour le bétail dans certaines parties du monde. Un système génétique de contrôle de la population a été proposé, dont le même type a déjà réussi chez certains insectes. Dans ce but, il est nécessaire de comprendre la génétique élémentaire de l'insecte, et pour cela, de réaliser entre autres une cartographie chromosomique. Cet animal possède six paires de chromosomes et sa cartographie est en cours.

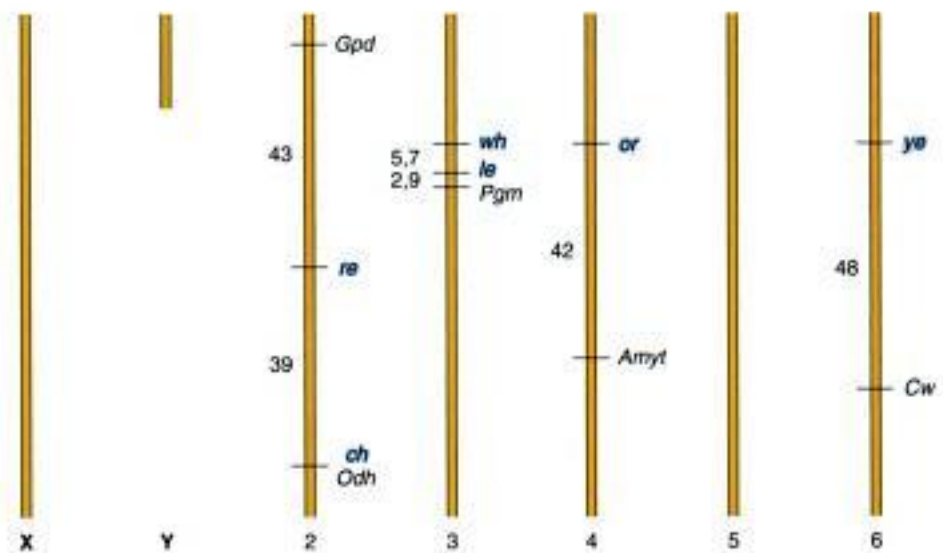
Une cartographie générale commence par la recherche et l'analyse du plus grand nombre de phénotypes variants possibles. Bien que les gènes contrôlant la plupart de ces phénotypes ne soient pas directement impliqués dans le projet de contrôle des mouches, ce sont des marqueurs nécessaires pour construire la carte chromosomique. Le stade adulte de cet insecte est une mouche et les généticiens ont découvert de nombreux phénotypes variants parmi les lucilles bouchères. Ils ont par exemple trouvé six couleurs différentes d'yeux, toutes distinctes de

la couleur marron que l'on observe chez les mouches de type sauvage, comme le montre la figure 5-13a. Ils ont également trouvé cinq phénotypes variants pour d'autres caractères. On a montré que 11 allèles mutants déterminaient les 11 phénotypes variants, chacun étant situé au niveau d'un locus autosomique différent. Les lignées pures de chaque phénotype ont été croisées les unes avec les autres pour produire des F_1 dihybrides et ont ensuite subi un croisement-test ou une autofécondation selon les méthodes décrites plus haut. L'analyse de ces



(a)

Figure 5-13 (a) Adulte de type sauvage de la lucille bouchère à côté de six mouches dont la couleur des yeux est déterminée par des allèles situés au niveau de six locus autosomiques différents. (b) Cartes de liaison génétique des six locus de la couleur des yeux (sur fond bleu) et de cinq autres locus chez la lucille bouchère. Les nombres indiqués entre les locus correspondent aux fréquences de recombinaison (D. B. Taylor, USDA).



(b)

mutants a révélé jusqu'ici quatre groupes de liaison présentés dans la figure 5-13b. Il est à noter que les locus *ye* et *cu* sont représentés provisoirement liés, bien que leur fréquence de recombinaison ne soit pas significativement différente de 50.

Les figures 5-14 et 5-15 présentent deux exemples de cartes de liaison : celle de la drosophile (*Drosophila*) et celle de la tomate. Le génome de la drosophile est l'un des génomes les plus intensivement cartographiés parmi tous les organismes génétiques modèles. La carte de la

figure 5-15 (pages 148-149) contient seulement une partie des locus connus. Les tomates ont également présenté un grand intérêt, tant sur le plan de la recherche génétique fondamentale que sur celui de la recherche appliquée et le génome de la tomate est l'un des mieux caractérisés chez les végétaux.

Les différentes parties de la figure 5-15 illustrent certains des niveaux de compréhension par lesquels la recherche doit passer pour produire une carte complète. Premièrement, bien que les chromosomes soient visibles

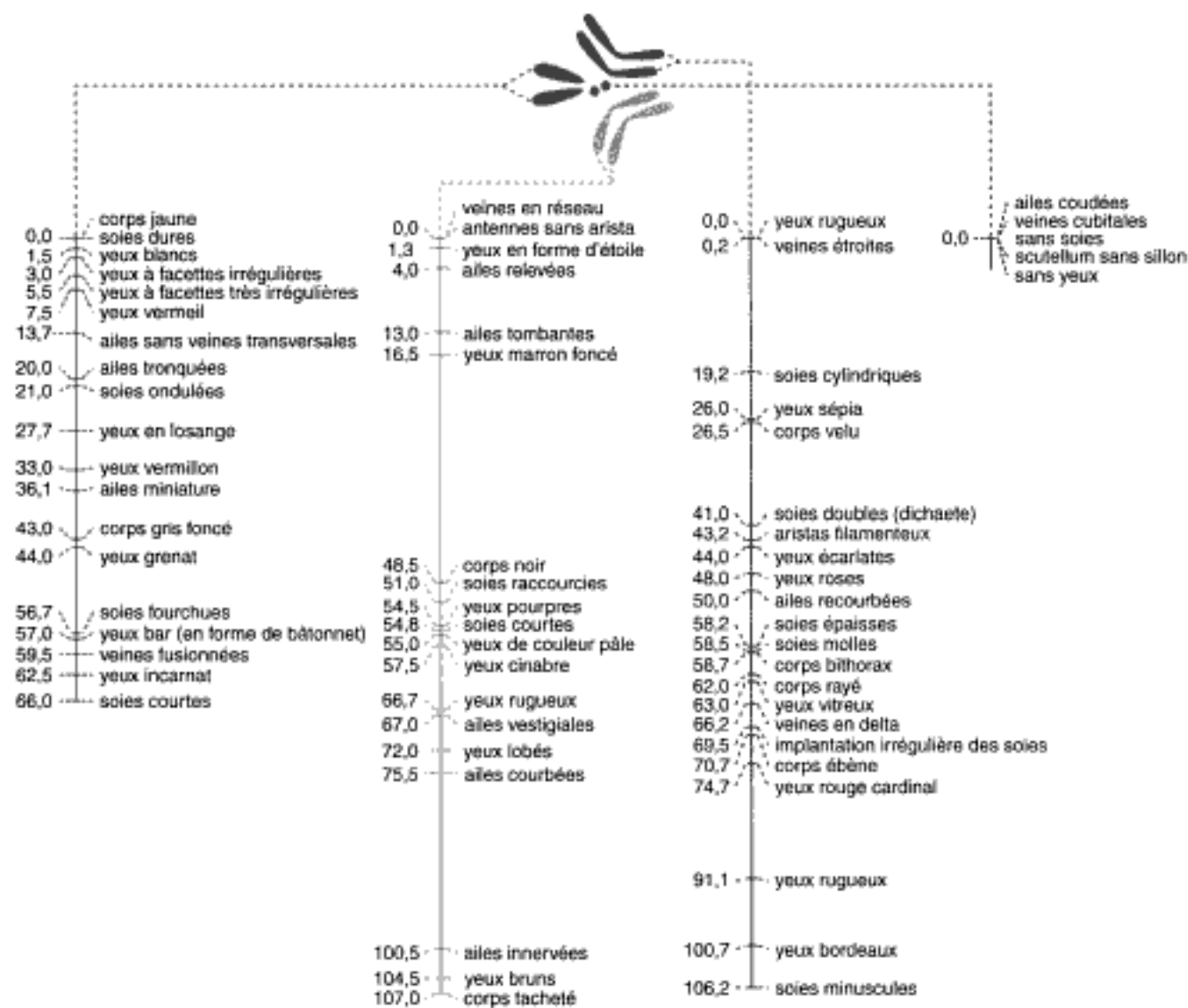


Figure 5-14 Carte génétique du génome de drosophile, montrant certains des locus et de quelle façon chaque groupe de liaison correspond à une paire de chromosomes. Les valeurs sont données en unités génétiques, mesurées à partir du gène le plus proche d'une extrémité. Les valeurs les plus importantes sont calculées par la somme d'intervalles plus courts car la fréquence de recombinaison de n'importe quelle paire de locus ne peut dépasser 50 pour cent. (D'après E. W. Sinnott, L. C. Dunn et T. Dobzhansky, *Principles of Genetics*, 5^e édition. Copyright © 1962 par McGraw-Hill).

Hidden page

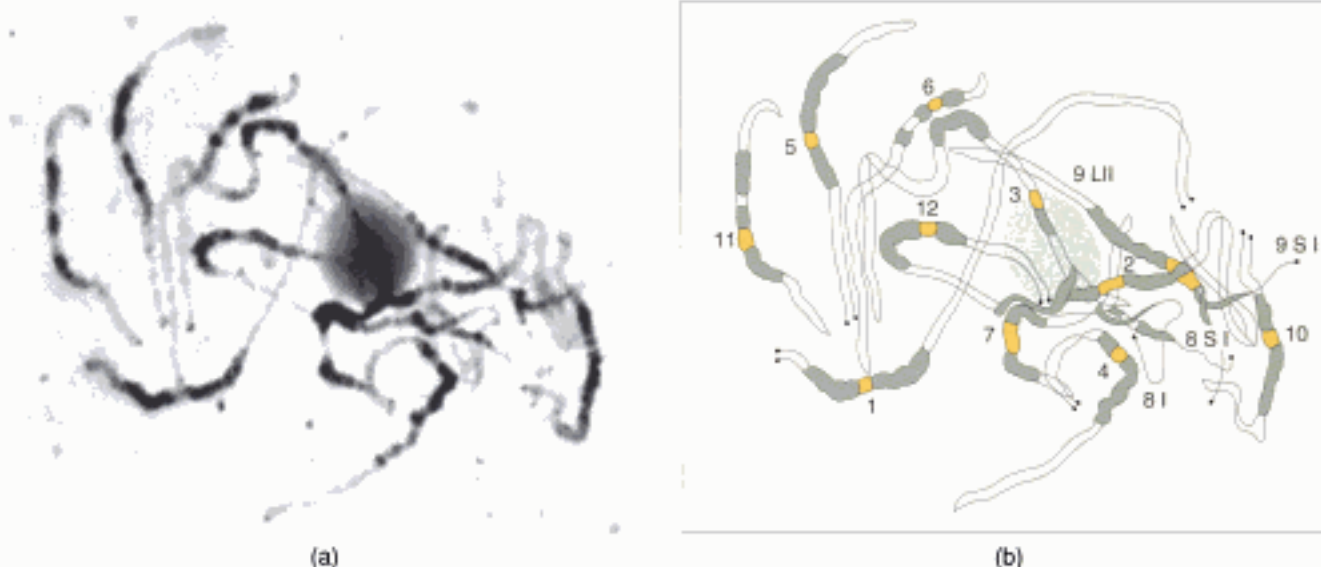


Figure 5-15 Cartographie des chromosomes de tomates. (a) Microphotographie d'une prophase méiotique I (pachytène) dans les anthères, montrant les 12 paires de chromosomes telles qu'on les voit sous microscope. (b) Le système actuel de numérotation des chromosomes. Les centromères sont en orange, et les régions adjacentes, à la coloration dense (hétérochromatine) sont en vert. (c) (page 149) Une carte de liaison génétique réalisée en 1952, montrant les groupes de liaison connus à cette époque. À côté de chaque locus se trouve un dessin du phénotype variant qui a permis d'identifier ce locus génétique (à droite) et un dessin du phénotype normal correspondant (à gauche). Les distances génétiques entre les locus sont indiquées en unités génétiques (Parties a et b d'après C. M. Rick, « The tomato », Copyright © 1978 par Scientific American Inc. Tous droits réservés. Partie c d'après L. A. Butler).

Étudions un groupe particulier de données de liaison génétique en utilisant l'analyse par le test du χ^2 . Supposons que nous ayons réalisé un croisement entre des parents de lignées pures, de génotypes A/A , B/B et a/a , b/b et que nous ayons obtenu un dihybride A/a , B/b auquel nous avons fait subir un croisement-test avec a/a , b/b . Les 500 descendants obtenus se répartissent de la façon suivante (écrits sous la forme des gamètes du dihybride) :

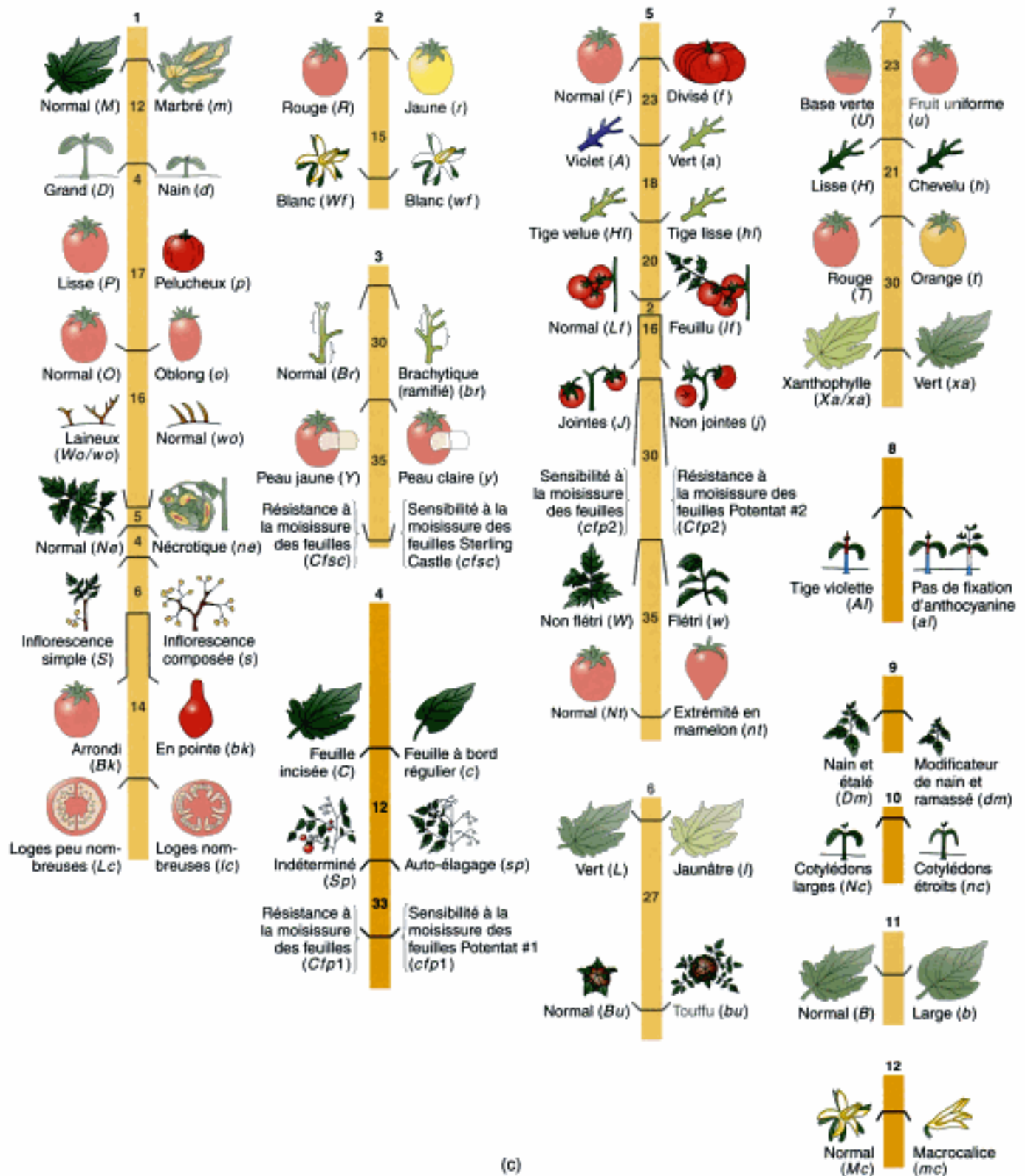
140	A, B parental
135	a, b parental
110	A, b recombinant
115	a, B recombinant
500	Total

D'après ces données, nous savons que la fréquence de recombinaison est de $225/500 = 45$ pour cent. À première vue, il semble que l'on se trouve dans un cas de liaison génétique, car la FR est inférieure aux 50 pour cent attendus pour un assortiment indépendant. Cependant, il est possible que les deux classes de recombinants soient inférieures à 50 pour cent seulement par hasard. Il nous faut donc réaliser un test du χ^2 . La statistique du χ^2 est toujours calculée à partir de nombres réels et non de pourcentages, proportions ou fractions. La taille de l'échantillon est donc très importante dans le test du χ^2 , comme pour la plupart des études de phénomènes aléatoires. Les

échantillons à étudier contiennent généralement plusieurs classes. La lettre O est généralement utilisée pour représenter le nombre observé dans une classe et la lettre A , le nombre attendu pour la même classe, évalué d'après l'hypothèse. La formule générale qui permet de calculer le χ^2 est la suivante :

$$\chi^2 = \text{total de } \frac{(O - A)^2}{A} \text{ pour toutes les classes}$$

Le problème est alors de calculer les résultats attendus A pour chaque classe. L'hypothèse qui doit être testée ici est celle d'un assortiment indépendant de deux locus, car l'hypothèse de liaison génétique ne peut fournir précisément les proportions attendues. Nous pouvons dire que si les paires d'allèles du dihybride subissent un assortiment aléatoire (c'est-à-dire, s'il n'y a pas de liaison génétique), il devrait y avoir un rapport 1 : 1 : 1 : 1 des types de gamètes. Il semble donc raisonnable d'utiliser 1/4 comme la proportion attendue de chaque classe de gamètes. Cependant, cette supposition est valable si la viabilité de tous les génotypes est égale, ce qui n'est pas nécessairement le cas. Un meilleur test consiste à savoir si oui ou non, les rapports des deux allèles sont combinés au hasard dans les produits de la méiose, car ceci ne correspond à aucune hypothèse *a priori* sur leur viabilité égale. Suivant cette approche, les classes phénotypiques sont écrites sous la forme d'un **tableau de contingence**



(c)

comme le montre le tableau 5-1. Le tableau est simplement une représentation graphique de la combinaison des rapports des allèles des deux locus. L'analyse par χ^2 teste si oui ou non les nombres d'individus dans les quatre classes phénotypiques de descendants sont simplement les deux rapports des allèles combinés au hasard, comme attendu d'après l'hypothèse d'assortiment indépendant.

TABLEAU 5-1

Tableau de contingence
pour les gènes *A/a* et *B/b*

	Ségrégation <i>B/b</i>		Total
	<i>B</i>	<i>b</i>	
	<i>A.B</i>	<i>A.b</i>	
<i>A</i>	$O = 140$ $E = 255 \times \frac{250}{500}$ $= 127,5$	$O = 110$ $E = 245 \times \frac{250}{500}$ $= 122,5$	250
<i>a</i>	$O = 115$ $E = 255 \times \frac{250}{500}$ $= 127,5$	$O = 135$ $E = 245 \times \frac{250}{500}$ $= 122,5$	250
Total	255	245	

Les valeurs attendues sont calculées simplement en subdivisant les totaux des colonnes proportionnellement aux totaux des lignes. La valeur de χ^2 se calcule de la façon suivante :

Classe	<i>O</i>	<i>A</i>	$\frac{(O-A)^2}{A}$
<i>A B</i>	140	127,5	1,23
<i>a b</i>	135	122,5	1,28
<i>A b</i>	110	122,5	1,28
<i>a B</i>	115	127,5	1,23
Total = $\chi^2 = 5,02$			

La valeur de χ^2 obtenue (5,02) est convertie en une valeur de probabilité en utilisant la table du χ^2 (tableau 5-2). Pour cela, il nous faut calculer le nombre de degrés de liberté (dl) utilisés dans le calcul du χ^2 , qui est le nombre d'écarts indépendants, $O - A$, mais tous les écarts dans le tableau de contingence se révèlent égaux en grandeur (12,5). Ceci s'explique par le fait que les nombres totaux attendus dans chaque ligne ou colonne doivent être égaux au nombre total observé dans chaque ligne et chaque colonne. Il y a donc un seul degré de liberté. En général, dans les tableaux de contingence, les degrés de liberté sont égaux au nombre de classes représentées dans les lignes moins 1, multiplié par le nombre de classes représentées dans les colonnes moins 1. Dans notre exemple :

$$dl = (2-1) \times (2-1) = 1$$

TABLEAU 5-2

Valeurs critiques de la distribution du χ^2

dl	p	0,995	0,975	0,9	0,5	0,1	0,05	0,025	0,01	0,005	dl
1		,000	,000	0,016	0,455	2,706	3,841	5,024	6,635	7,879	1
2		0,010	0,051	0,211	1,386	4,605	5,991	7,378	9,210	10,597	2
3		0,072	0,216	0,584	2,366	6,251	7,815	9,348	11,345	12,838	3
4		0,207	0,484	1,064	3,357	7,779	9,488	11,143	13,277	14,860	4
5		0,412	0,831	1,610	4,351	9,236	11,070	12,832	15,086	16,750	5
6		0,676	1,237	2,204	5,348	10,645	12,592	14,449	16,812	18,548	6
7		0,989	1,690	2,833	6,346	12,017	14,067	16,013	18,475	20,278	7
8		1,344	2,180	3,490	7,344	13,362	15,507	17,535	20,090	21,955	8
9		1,735	2,700	4,168	8,343	14,684	16,919	19,023	21,666	23,589	9
10		2,156	3,247	4,865	9,342	15,987	18,307	20,483	23,209	25,188	10
11		2,603	3,816	5,578	10,341	17,275	19,675	21,920	24,725	26,757	11
12		3,074	4,404	6,304	11,340	18,549	21,026	23,337	26,217	28,300	12
13		3,565	5,009	7,042	12,340	19,812	22,362	24,736	27,688	29,819	13
14		4,075	5,629	7,790	13,339	21,064	23,685	26,119	29,141	31,319	14
15		4,601	6,262	8,547	14,339	22,307	24,996	27,488	30,578	32,801	15

Donc, en utilisant le tableau 5-2, nous regardons le long de la ligne 1 degré de liberté jusqu'à ce que nous trouvions notre valeur de χ^2 de 5,02. Nous voyons que la probabilité d'obtenir un écart par rapport à la valeur du résultat que l'on attend (ou à une valeur supérieure) par le seul hasard est de 0,025 (2,5 pour cent). Puisque cette probabilité est inférieure à 5 pour cent, l'hypothèse d'un assortiment indépendant doit être rejetée. En ayant rejeté l'hypothèse d'une absence de liaison génétique, nous en déduisons donc que les locus doivent être liés.

MESSAGE

Le test du χ^2 est utilisé pour comparer des nombres observés dans une classe, aux résultats attendus d'après une hypothèse. Le test donne la probabilité d'obtenir par hasard un écart au moins aussi élevé que celui observé, en supposant que l'hypothèse est juste.

LE CROSSING-OVER MITOTIQUE

Les généticiens furent surpris en découvrant que des crossing-over pouvaient également se produire lors de la mitose. Cela doit sans doute avoir lieu lorsque des segments de chromosomes homologues sont accidentellement appariés dans des cellules asexuées (somatiques) telles que les cellules du corps. Les crossing-over mitotiques sont rares la plupart du temps, mais chez certains organismes, ils peuvent être fréquents — par exemple quelques champignons qui n'ont pas de cycle sexuel utilisent les crossing-over mitotiques comme source de variation génétique. De plus, dans l'étude des cancers humains, on pense que le crossing-over mitotique est important car il permettrait l'expression de mutations récessives induisant un cancer.

Les crossing-over mitotiques se produisent uniquement dans les cellules diploïdes, telles que les cellules somatiques des organismes diploïdes. (Pour étudier les crossing-over mitotiques chez un organisme haploïde, il faut créer artificiellement des cellules diploïdes).

Regardons un exemple tiré de la drosophile, impliquant des gènes de la couleur du corps liés génétiquement (j^+ = marron, j = jaune) et de la forme des soies (sn^+ = droit, sn = ondulé [singé en anglais]). Une mouche qui est un hybride *trans* pour ces allèles est présentée dans la figure 5-16. Un dihybride *trans* devrait exprimer les phénotypes dominants pour les deux gènes dans les cellules de tout le corps. Ce sera évidemment le cas pour la plupart des cellules. Cependant, si un crossing-over mitotique se produit, ceci peut conduire à la formation de combinaisons alléliques homozygotes pour tous les gènes situés sur le bras chromosomique *distal* par rapport au crossing-over (c'est-à-dire la portion du bras qui se trouve au-delà du crossing-over en s'éloignant du centromère). Les allèles récessifs exprimeront donc de façon inattendue ce phénotype. Puisque les cellules produites par un crossing-over mitotique engendreront un groupe de cellules filles qui auront tendance à rester ensemble au cours

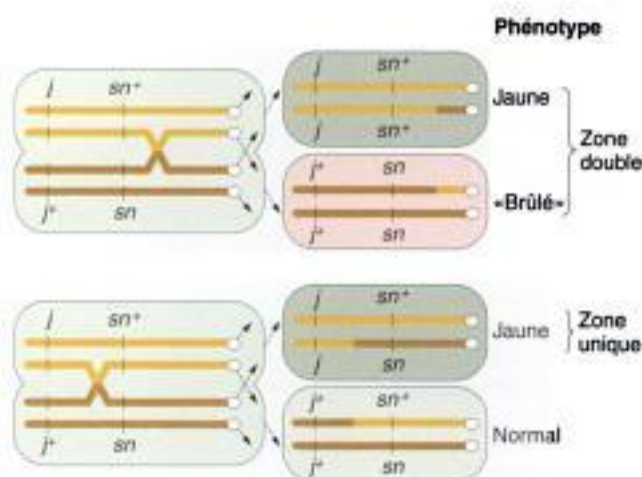


Figure 5-16 Un crossing-over mitotique peut conduire à l'homozygotie de locus distaux.

du développement, les phénotypes récessifs s'exprimeront sous la forme de taches ou de points sur le corps. Dans l'exemple de la figure 5-16, deux taches jumelles de phénotypes récessifs sont produites par un crossing-over entre le centromère et le locus le plus proche et une tache jaune unique est produite par un crossing-over entre les locus (figure 5-17).

MESSAGE

Un crossing-over mitotique engendre, dans un groupe de cellules, une homozygotie des allèles de locus hétérozygotes, distaux par rapport au crossing-over.

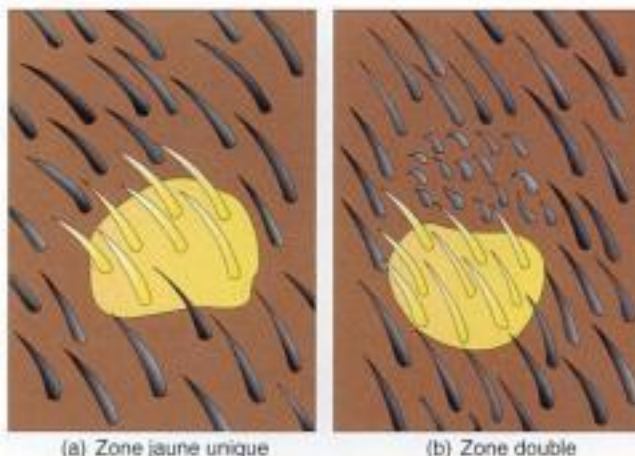


Figure 5-17 Secteurs inattendus de phénotypes de la surface du corps, produits par des crossing-over mitotiques chez la drosophile. La mouche a un génotype $j sn^+ / j^+ sn$ (où j représente un corps jaune et sn des soies « brûlées »).

Hidden page

structure instable impliquant un mésappariement de l'ADN — peut-être une paire G•T — et on connaît des enzymes qui reconnaissent de tels mésappariements. Elles coupent l'un des brins mal appariés et le remplacent par un brin contenant un nucléotide dont la liaison hydrogène est correcte (figure 5-19). Ce type de réparation d'un mésappariement dans une structure où il n'y a pas eu de déplacement de la fourche, donnerait soit un rapport 6 : 2, ou renverrait à un rapport 4 : 4 normal. Dans les ADN hétérologues résultant du déplacement de la fourche, la réparation peut produire n'importe lequel des autres rapports aberrants. Dans ces régions, s'il n'y a pas de réparation, on obtient le rapport 3 : 1 : 1 : 3.

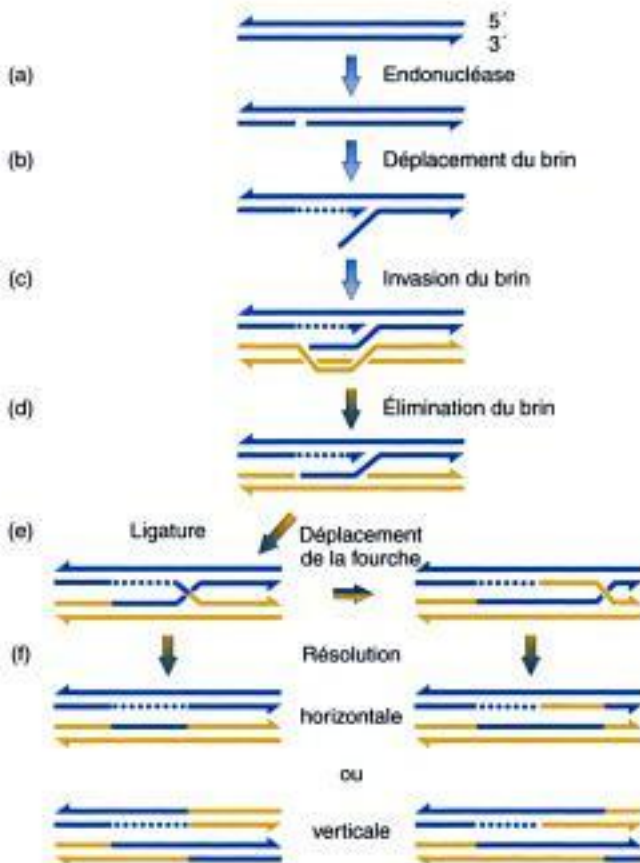


Figure 5-18 Le modèle de l'ADN double-brin hétérologue de Meselson-Radding. (a) Un double-brin est coupé au niveau d'un brin. (b) L'ADN polymérase déplace un brin. (c) Le simple-brin résultant déplace son équivalent chez l'homologue. (d) Ce brin déplacé est digéré enzymatiquement. (e) Une ligature complète la formation de la jonction de Holliday. Celle-ci est génétiquement asymétrique car seul un des deux doubles-brins possède une région dans laquelle peut se former un ADN double-brin hétérologue. Si la jonction migre, un ADN double-brin hétérologue peut apparaître sur les deux doubles-brins. (f) La résolution de la jonction se produit comme dans le modèle de Holliday. (D'après F. W. Stahl, « The Holliday junction on its thirtieth anniversary », *Genetics*, 138, 1994, 241-246).

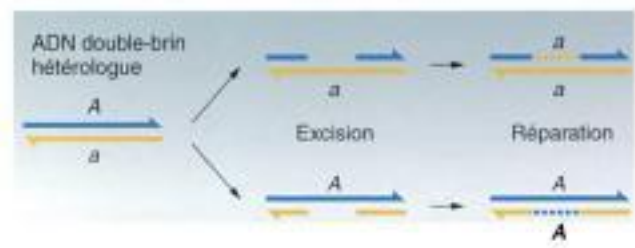


Figure 5-19 Réparation de nucléotides mal appariés dans un ADN double-brin hétérologue, qui redonne soit A, soit a. L'ADN double brin hétérologue contenant des paires de bases mal appariées est instable. Les nucléotides mal appariés peuvent entraîner une déformation de l'ADN à ces endroits, qui seront reconnus par un système de réparation.

Le modèle de Meselson-Radding est seulement l'un des nombreux modèles basés sur la formation d'ADN hétérologue. Le mécanisme de recombinaison dans ce modèle est initié par une coupure simple-brin dans l'ADN. Dans un autre modèle, le mécanisme de recombinaison est initié par une coupure double-brin et ce modèle est soutenu par des études réalisées sur la recombinaison des levures. Toutefois, il n'y a pas un modèle applicable à tous les Eucaryotes.

MESSAGE

L'observation des rares cas de paires d'ascospores non identiques dans des octades de champignons suggère que des molécules d'ADN hétérologue peuvent se former et sont à la base des crossing-over. L'importance de la région d'ADN hétérologue et la réparation des mésappariements dans cet ADN peut expliquer tous les rapports aberrants observés.

LA RECOMBINAISON À L'INTÉRIEUR D'UN GÈNE

La méiose, chez un diploïde hétérozygote pour deux allèles différents du même gène (disons, a^1/a^2) aboutira à des produits dont la moitié seront a^1 et l'autre moitié a^2 , d'après la loi de ségrégation des allèles. Cependant, il est probable que ces deux allèles mutants aient un site modifié en deux endroits différents du gène et, si c'est le cas, il y a une faible probabilité pour que la recombinaison se produise entre les deux sites. Les produits seront une chromatide de type sauvage et une chromatide mutante avec deux sites mutants :

Hidden page

PROBLÈME RÉSOLU

1. Considérez trois pois lisses jaunes, A, B et C. Chacun a poussé et est devenu une plante, puis a été croisé avec une plante issue d'un pois vert ridé. 100 pois exactement ont été obtenus à la suite de chaque croisement et ont été triés suivant les classes phénotypiques indiquées ci-dessous.

A : 51	jaune, lisse
49	vert, lisse
B : 100	jaune, lisse

C : 24	jaune, lisse
26	jaune, ridé
25	vert, lisse
25	vert, ridé

Quels étaient les génotypes de A, B et C ? (Utilisez les symboles des gènes de votre choix : assurez-vous de définir chacun d'eux.)

SOLUTION

Remarquez que chacun des croisements est

Jaune, lisse \times vert, ridé
 \downarrow
 Descendants

A, B et C ayant été croisés avec la même plante, toutes les différences entre les trois populations de descendants doivent être attribuées à des différences entre les génotypes sous-jacents de A, B et C.

Que pouvons-nous déduire de ces données ? Qu'en est-il de la dominance ? Le croisement clé pour déduire la dominance est B. Ici, le mode de transmission est

Jaune, lisse \times vert, ridé
 \downarrow
 Tous jaune, lisse

Jaune et lisse doivent donc être les phénotypes dominants, car par définition, c'est le phénotype d'un hétérozygote qui définit la dominance. Maintenant que nous savons que le parent vert, ridé utilisé dans chaque croisement était entièrement récessif, nous sommes dans une situation très confortable, car cela signifie que chaque croisement est un croisement-test, ce qui est généralement le type de croisement le plus informatif.

Si nous revenons à la descendance de A, nous observons un rapport 1 : 1 entre le jaune et le vert. Ceci est une démonstration de la première loi de Mendel (ségrégation égale des allèles) et montre que, pour le caractère de la couleur, le croisement a dû être hétérozygote \times homozygote récessif. En posant J = jaune et j = vert, nous avons

$Jj \times jj$
 \downarrow
 $\frac{1}{2} Jj$ (jaune)
 et
 $\frac{1}{2} jj$ (vert)

Pour le caractère de la forme, comme tous les descendants sont lisses, le croisement a dû être homozygote dominant \times homozygote récessif. En posant R = lisse et r = ridé, nous avons

$R/R \times r/r$
 \downarrow
 R/r (lisse)

En combinant les deux caractères, nous obtenons

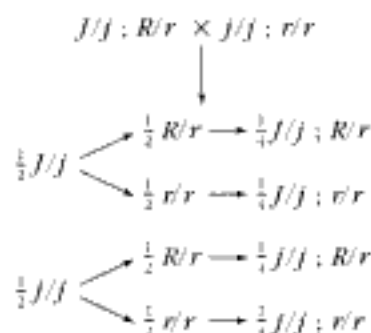
$Jj ; R/R \times jj ; r/r$
 \downarrow
 $\frac{1}{2} Jj ; R/r$
 $\frac{1}{2} jj ; R/r$

Le croisement B devient à présent clair comme de l'eau de roche et doit avoir été

$JJ ; R/R \times jj ; r/r$
 \downarrow
 $Jj ; R/r$

car toute hétérozygotie dans le pois B aurait conduit à plusieurs phénotypes de descendants, et non à un seul.

Qu'en est-il de C ? Ici nous avons un rapport de 50 jaune : 50 vert (1 : 1) et un rapport de 49 lisse : 51 ridé (également 1 : 1). Les deux gènes dans le pois C devaient par conséquent être hétérozygotes, et le croisement C était



ce qui est une bonne démonstration de la deuxième loi de Mendel (l'assortiment indépendant de gènes différents).

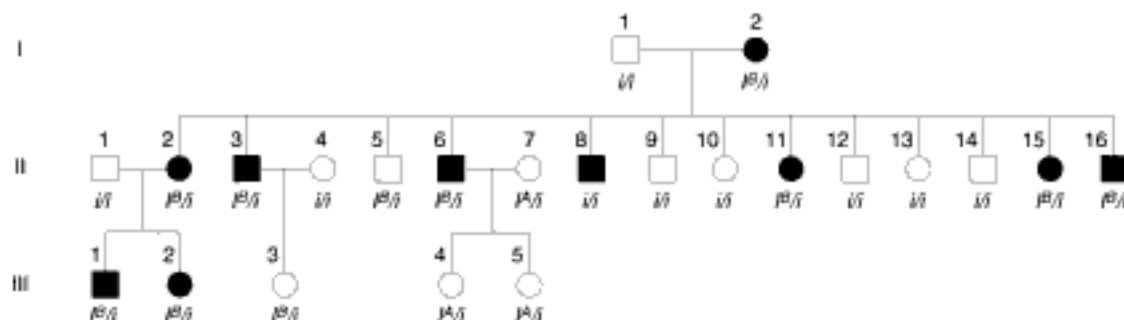
Comment un généticien aurait-il analysé ces croisements ? Globalement de la même façon que nous, mais avec un peu moins d'étapes intermédiaires. Sans doute un peu comme ça : - jaune et lisse dominants ; ségrégation d'un gène unique dans A ; B homozygote dominant ; ségrégation de deux gènes indépendants dans C. •

PROBLÈME RÉSOLU

2. Un arbre généalogique humain montre des personnes affectées du syndrome rare d'onychartrose (ongles et rotules mal formés) et indique également le génotype du groupe sanguin de chaque individu. (Remarque : les types sanguins ABO sont déterminés par trois allèles I^A , I^B et i . $I^A I^A$ détermine le groupe A, $I^B I^B$ le groupe B et $i i$ le groupe O. L'hétérozygote $I^A I^B$ donne le groupe AB.) Les deux locus concernés sont autosomiques. Pour répondre aux questions suivantes, étudiez l'arbre généalogique ci-dessous.

- Le syndrome d'onychartrose est-il un phénotype dominant ou récessif ? Donnez des raisons pour étayer votre réponse.
- Y a-t-il des preuves de liaison génétique entre le gène d'onychartrose et le gène de type sanguin ABO, comme le laisserait penser l'arbre généalogique ? Pourquoi oui ou pourquoi non ?

- S'il y a des preuves de liaison génétique, représentez les allèles portés par les homologues correspondants des grands-parents. S'il n'y a pas de preuve de liaison génétique, représentez les allèles sur les deux paires d'homologues.
- Selon votre modèle, quels descendants représentent les recombinants ?
- Quelle est la meilleure estimation de la FR ?
- Si l'homme III-1 se marie avec une femme normale de type sanguin O, quelle est la probabilité que leur premier enfant soit de type sanguin B et qu'il soit affecté du syndrome d'onychartrose ?



Hidden page

trihybrides, etc. ? Faites de même pour les phénotypes.

- d. Mendel prédit qu'à toutes les classes phénotypiques du carré de Punnett sauf une, devraient correspondre plusieurs génotypes différents. En particulier, il réalisa de nombreux croisements pour identifier les génotypes sous-jacents au phénotype lisse, jaune. Présentez deux façons différentes d'identifier les différents génotypes sous-jacents au phénotype lisse, jaune. (Rappelez-vous que tous les pois lisses, jaunes se ressemblent).

4. Le croisement-test suivant a été effectué pour une plante :

$$\begin{array}{cc} A & B \\ \hline a & b \end{array} \times \begin{array}{cc} a & b \\ \hline a & b \end{array}$$

Si les deux locus sont éloignés de 10 u.g., quelle proportion de descendants sera AB/ab ?

5. Le locus A et le locus D sont si étroitement liés qu'on n'observe jamais de recombinaison entre eux. Si Ad/Ad est croisé avec aD/aD et que les membres de la F_1 sont croisés les uns avec les autres, quels phénotypes observera-t-on dans la F_2 et dans quelles proportions ?
6. Une souche de *Neurospora* avec le génotype HJ est croisée avec une souche de génotype $h.i$. La moitié de la descendance est HJ et la moitié est $h.i$. Expliquez comment cela est possible.
7. Un animal femelle de génotype $A/a, B/b$ est croisé avec un mâle double récessif $a/a, b/b$. Leurs descendants comprennent 442 $A/a, B/b$, 458 $a/a, b/b$, 46 $A/a, b/b$ et 54 $a/a, B/b$. Expliquez ces proportions et représentez les chromosomes du parent dihybride en indiquant les positions des gènes avec leurs allèles.
8. Si $A/A, b/b$ est croisé avec $a/a, B/B$ et que la F_1 subit un croisement-test, quel pourcentage des descendants du croisement seront $a/a, b/b$ si les deux gènes (a) ne sont pas liés, (b) sont complètement liés (pas de crossing-over du tout), (c) sont distants de 12 unités génétiques, (d) sont distants de 24 unités génétiques ?
9. Une drosophile de génotype $B R/b r$ subit un croisement-test avec une mouche de génotype $b r/b r$. Dans 84 pour cent des méioses, il n'y a pas de chiasma entre les gènes liés ; dans 16 pour cent des méioses, il y a un chiasma entre les gènes. La proportion des descendants $B r/b r$ sera-t-elle (a) de 50 pour cent, (b) de 4 pour cent, (c) de 84 pour cent, (d) de 25 pour cent, (e) de 16 pour cent ?

10. Chez les chiens, la couleur noire du pelage est dominante sur le phénotype albinos et les poils courts sont dominants sur les poils longs. En supposant que ces effets sont causés par deux gènes dont l'assortiment est indépendant écrivez les génotypes des parents de chacun des croisements présentés ci-dessous, dans lesquels N et A désignent les phénotypes noir et albinos et où C et L correspondent aux phénotypes poils courts et poils longs.

Phénotypes parentaux	Nombre de descendants			
	N, C	N, L	A, C	A, L
a. $N, C \times N, C$	89	31	29	11
b. $N, C \times N, L$	18	19	0	0
c. $N, C \times A, C$	20	0	21	0
d. $A, C \times A, C$	0	0	28	9
e. $N, L \times N, L$	0	32	0	10
f. $N, C \times N, C$	46	16	0	0
g. $N, C \times N, L$	30	31	9	11

Utilisez les symboles N et n pour les allèles noir et albinos de la couleur du pelage et C et c pour les allèles des poils courts et des poils longs respectivement. Faites l'hypothèse de l'homozygotie, à moins qu'il n'y ait des preuves d'une autre situation.

(Problème 10 réimprimé avec la permission de Macmillan Publishing Co., Inc., d'après *Genetics* de M. Strickberger. Copyright © 1968 par Monroe W. Strickberger.)

11. Chez les tomates, deux allèles d'un gène déterminent les phénotypes violet (Vl) et vert (Ve) des tiges et deux allèles d'un gène distinct, indépendant, déterminent les phénotypes - incisé - (I) et - à bord régulier - (Br) des feuilles. Les résultats de cinq croisements de plants de tomates de divers phénotypes sont présentés ci-dessous :

Croisement	Phénotypes parentaux	Nombre de descendants			
		Vl, I	Vl, Br	G, I	G, Br
1	$Vl, I \times G, I$	321	101	310	107
2	$Vl, I \times Vl, Br$	219	207	64	71
3	$Vl, I \times G, I$	722	231	0	0
4	$Vl, I \times G, Br$	404	0	387	0
5	$Br \times G, I$	70	91	86	77

- a. Déterminez les phénotypes dominants.
- b. Quels sont les génotypes les plus probables pour les parents de chaque croisement ?

(Problème 11 modifié d'après A. M. Srb, R. D. Owen et R. S. Edgar, *General Genetics*, 2^e éd. Copyright © 1965 par W. H. Freeman and Company.)

12. Nous avons vu jusqu'à présent essentiellement des situations d'assortiment indépendant de deux gènes seulement, mais les mêmes principes sont valables pour plus de deux gènes. Considérons le croisement :

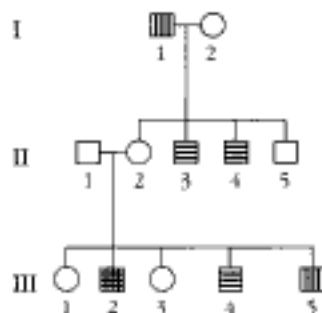
$$A/a; B/b; C/c; D/d; E/e \times a/a; B/b; c/c; D/d; e/e$$

- Quelle proportion de descendants ressembleront *phénotypiquement* (1) à leur premier parent, (2) à leur deuxième parent, (3) à l'un ou l'autre de leurs parents et (4) à aucun de leurs parents ?
 - Quelle proportion de descendants seront *génotypiquement* identiques (1) à leur premier parent, (2) à leur deuxième parent, (3) à l'un ou l'autre de leurs parents et (4) à aucun de leurs parents ?
13. Après plusieurs croisements de type général $A/A; B/B \times a/a; b/b$, les individus de la F_1 de type $A/a; B/b$ subissent un croisement-test avec $a/a; b/b$. Les résultats sont indiqués ci-dessous :

Croisement-test de la F_1 issue du croisement	Descendants du croisement-test			
	$A/a; B/b$	$a/a; b/b$	$A/a; b/b$	$a/a; B/b$
1	310	315	287	288
2	36	38	23	23
3	360	380	230	230
4	74	72	50	44

Pour chaque groupe de descendants, utilisez le test du χ^2 pour déterminer s'il y a liaison génétique.

14. Dans l'arbre généalogique suivant, les lignes verticales correspondent à une cécité protéranope (cécité au rouge ; individu dyschromate rouge moins) et les lignes horizontales à une cécité deutéranope (cécité au vert ; individu dyschromate vert moins). Il s'agit de maladies distinctes entraînant de mauvaises perceptions différentes des couleurs ; chacune est déterminée par un gène distinct.



- L'arbre généalogique montre-t-il des preuves de liaison entre les gènes ?
- S'il y a liaison génétique, l'arbre généalogique montre-t-il des preuves de crossing-over ?

Expliquez vos deux réponses à l'aide d'un schéma.

- Pouvez-vous calculer une valeur de recombinaison entre ces gènes ? Cette recombinaison provient-elle d'un assortiment indépendant ou de crossing-over ?

15. Un croisement-test à trois points a été effectué avec du maïs. Les résultats et une analyse de recombinaison partielle sont présentés dans le tableau suivant (v = feuilles violettes, $+$ = vertes ; r = semis résistants au virus, $+$ = sensible, m = bande marron sur la graine, $+$ = dépourvu de bande). Examinez le tableau et répondez aux questions ci-dessous.

P	$+/+++/+++/+ \times v/v \cdot r/r \cdot m/m$
Gamètes	$+++ \quad v \cdot r \cdot m$
F_1	$+/v \cdot +/r \cdot +/m \times v/v \cdot r/r \cdot m/m$ (souche-test)

Classe	Phénotypes des descendants	Gamètes de la F_1	Nombre	Recombinant pour		
				$v \cdot m$	$v \cdot r$	$r \cdot m$
1	ver sen dep	$+++$	3 210			
2	vio rés ban	$v \cdot r \cdot m$	3 222			
3	ver rés dep	$+r+$	1 024		R	R
4	vio sen ban	$v \cdot +m$	1 044		R	R
5	vio rés dep	$v \cdot r \cdot +$	690	R		R
6	ver sen ban	$+++m$	678	R		R
7	ver rés ban	$+r \cdot m$	72	R	R	
8	vio sen dep	$v \cdot ++$	60	R	R	
Total 10 000				1 500	2 200	3 436

- Déterminez les gènes qui sont liés.
- Dessinez une carte qui montre les distances en unités génétiques.
- Calculez l'interférence s'il y a lieu.

Décomposons le problème

- Faites des croquis du parent (P), de la F_1 et des souches-test de plants de maïs et utilisez des flèches pour montrer exactement comment vous réaliseriez cette expérience. Montrez à quel endroit les graines sont ramassées.
- Pourquoi tous les $+$ se ressemblent-ils, même lorsqu'ils ont des gènes différents ? Pourquoi cela n'entraîne-t-il pas de confusion ?
- Comment un phénotype peut-il être violet et marron (par exemple) en même temps ?
- Est-il important que les gènes soient écrits dans l'ordre $v \cdot r \cdot m$ dans le problème ?

- e. Qu'est-ce qu'une souche-test et pourquoi est-elle utilisée dans cette analyse ?
- f. Que représente la colonne marquée « phénotypes des descendants » ? Dans la classe 1 par exemple, dites exactement ce que signifie « ver sen dep ».
- g. Que représente la ligne marquée « gamètes » et en quoi est-elle différente de la colonne marquée « gamètes de la F_1 » ? En quoi la comparaison de ces deux types de gamètes a-t-elle un rapport avec la recombinaison ?
- h. Quelle méiose est le sujet d'étude principal ? Marquez-la sur votre dessin.
- i. Pourquoi les gamètes de la souche-test ne sont-ils pas indiqués ?
- j. Pourquoi y a-t-il seulement huit classes phénotypiques ? En manque-t-il ?
- k. Quelles classes (et dans quelles proportions) attendez-vous si tous les gènes se trouvent sur des chromosomes séparés ?
- l. À quoi correspondent les quatre types effectifs des classes (deux très grands, deux intermédiaires, deux intermédiaires et deux très petits) ?
- m. Que pouvez-vous dire sur l'ordre des gènes en examinant les classes phénotypiques et leur fréquence ?
- n. Quelle serait la distribution attendue des classes phénotypiques si seulement deux gènes étaient liés ?
- o. Que signifie le *point* dans un croisement-test à trois points ? Son usage implique-t-il une liaison génétique ? À quoi ressemblerait un croisement-test à quatre points ?
- p. Quelle est la définition de *recombinant* et comment est-elle appliquée ici ?
- q. Que signifient les colonnes « recombinant pour » ?
- r. Pourquoi y a-t-il seulement trois colonnes « recombinant pour » ?
- s. Que signifient les « R » et comment sont-ils déterminés ?
- t. Que signifient les totaux des colonnes ? Comment sont-ils utilisés ?
- u. Quel test utilise-t-on pour vérifier s'il y a liaison génétique ?
- v. Qu'est-ce qu'une unité génétique ? Est-ce la même chose qu'un centimorgan ?
- w. Dans un croisement-test à trois points tel que celui-là, pourquoi ne considère-t-on pas la F_1 et la souche-test comme parentales lorsqu'on calcule la recombinaison ? (Ce *sont* des parents dans un sens).
- x. Quelles est la formule de l'interférence ? Comment les fréquences « attendues » sont-elles calculées dans la formule du coefficient de coïncidence ?
- y. Pourquoi est-il écrit « s'il y a lieu » dans la partie c du problème ?
- z. Quel travail cela représente-t-il d'obtenir un tel nombre de descendants dans le maïs ? Combien de descendants sont approximativement représentés par épi de maïs ?
16. R. A. Emerson a croisé deux lignées pures de maïs et a obtenu une F_1 de phénotype sauvage qui était hétérozygote pour les allèles de trois gènes différents, qui déterminent des phénotypes récessifs : *an* détermine le phénotype mutant anthère, *br* brachytique et *f* : fin. Il a réalisé un croisement entre la F_1 et une souche-test qui était homozygote récessive pour les trois gènes. Il a obtenu les phénotypes suivants dans la descendance : 355 anthère, 339 brachytique, fin ; 88 entièrement de type sauvage ; 55 anthère, brachytique, fin ; 21 fin ; 17 anthère, brachytique ; 2 brachytique ; 2 anthère, fin.
- a. Quels étaient les génotypes des lignées parentales ?
- b. Dessinez une carte de liaison génétique pour les trois gènes (comportant les distances génétiques).
- c. Calculez la valeur d'interférence.
17. D'après les cinq groupes de données indiquées dans le tableau ci-dessous, déterminez l'ordre des gènes par simple examen — c'est-à-dire, sans calculer les valeurs de recombinaison. Les phénotypes récessifs sont symbolisés par des lettres minuscules et les phénotypes dominants par des plus.

Phénotypes observés dans un croisement-test à 3 points	Résultats				
	1	2	3	4	5
+++	317	1	30	40	305
++c	58	4	6	232	0
+b+	10	31	339	84	28
+bc	2	77	137	201	107
a++	0	77	142	194	124
a+c	21	31	291	77	30
ab+	72	4	3	235	1
abc	203	1	34	46	265

18. Chez la plante *Arabidopsis thaliana*, les locus de la longueur des cosses (L : long, l = court) et des poils des fruits (C = chevelu, c = lisse) sont liés génétiquement et distants de 16 unités génétiques sur le même chromosome. Les croisements suivants ont été effectués :

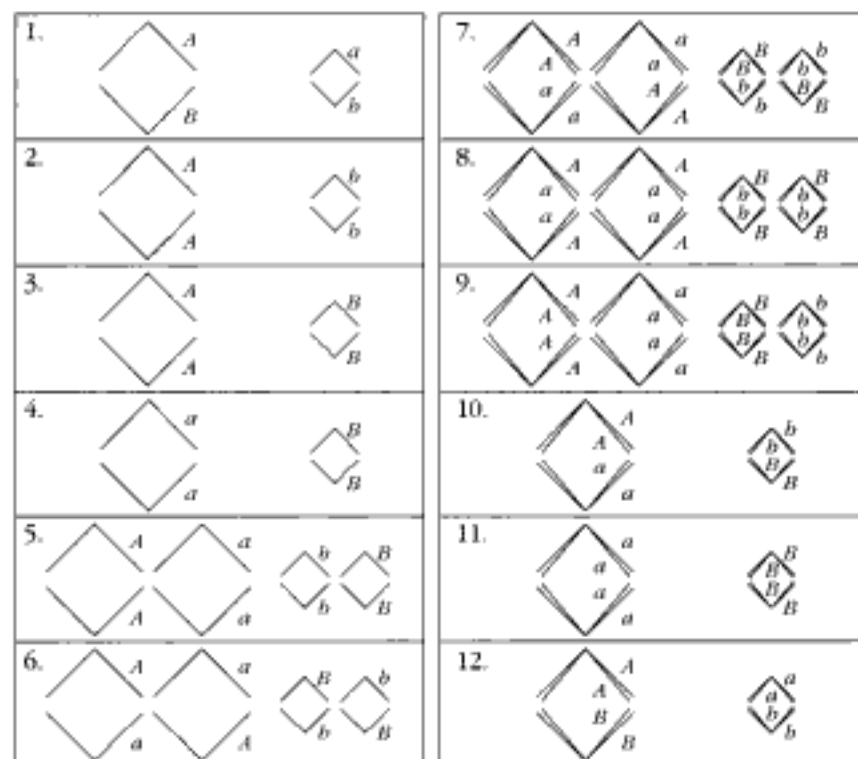
$$LCLC \times lclc \longrightarrow F_1$$

$$LclC \times lClC \longrightarrow F_1$$

Si les plantes de la F_1 obtenues ci-dessus sont croisées l'une par l'autre,

- Quelle proportion de descendants $lclc$ s'attend-on à obtenir ?
 - Quelle proportion de descendants $LclC$ s'attend-on à obtenir ?
19. Supposez que chez un champignon ascomycète, un ADN hétérologue se forme au niveau d'un locus hétérozygote A/a (consultez la figure 5-18). Supposez qu'un brin A envahisse une hélice a à la même fréquence qu'un brin a envahit une hélice A . Supposez qu'il n'y ait pas de déplacement de la fourche. Supposez également que les ADN hétérologues soient réparés en formant A , 80 pour cent du temps et a , 20 pour cent du temps. Parmi tous les cas de formation d'un ADN hétérologue à ce site, quelle proportion d'ADN hétérologues seront convertis en rapports aberrants 6 : 2, 2 : 6, 5 : 3 et 3 : 5 ?

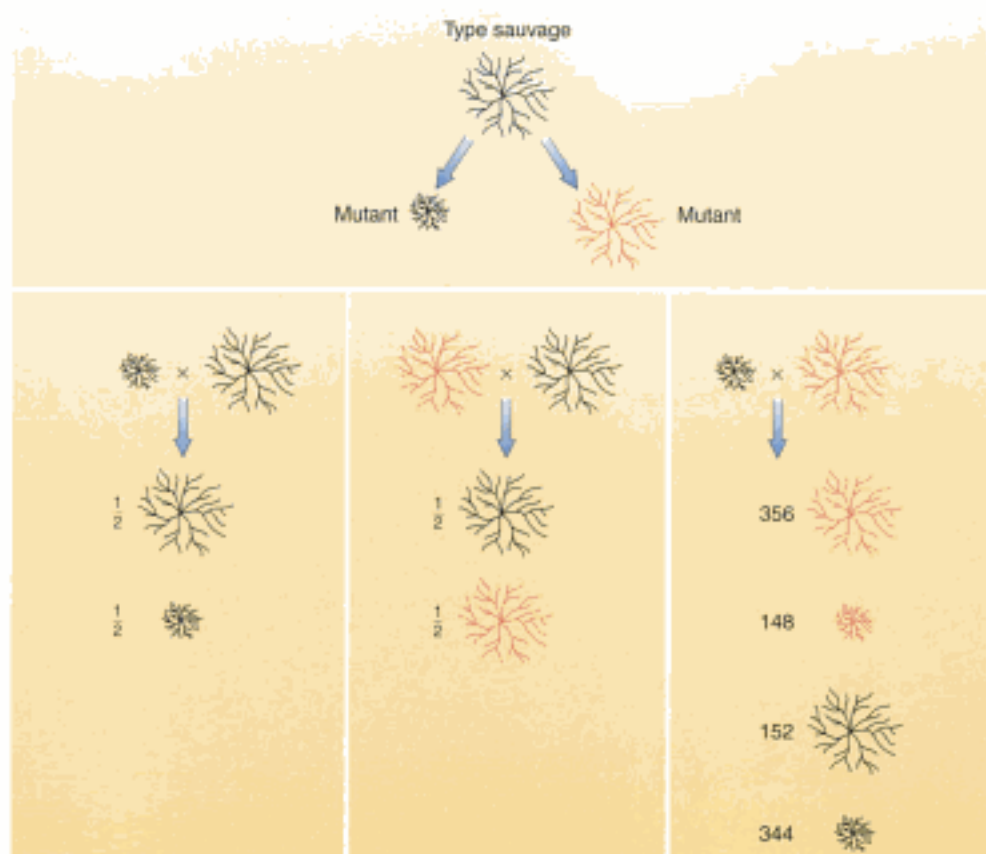
20. La plante *Haplopappus gracilis* est diploïde et $2n = 4$. Elle possède une paire de chromosomes courts et une paire de chromosomes longs. Les schémas ci-contre représentent des anaphases (stades auxquels les chromosomes sont tirés vers les pôles) de cellules individuelles en cours de méiose ou de mitose, chez une plante qui est génétiquement dihybride (A/a ; B/b) pour des gènes situés sur des chromosomes différents. Les traits représentent des chromosomes ou des chromatides et les pointes des V représentent les centromères. Dans chaque cas, déterminez si le schéma représente une cellule en méiose I, en méiose II ou en mitose. Si un schéma présente une situation impossible, dites-le.



21. Le gène *hist-1* de *Neurospora*, qui se trouve sur le bras droit du chromosome 1, est impliqué dans la voie métabolique de synthèse de l'histidine. Deux mutations du gène *hist-1* ont été obtenues à la suite d'expériences indépendantes. Ces deux mutations ont rendu les souches auxotrophes — c'est-à-dire qu'elles ne croissent pas sans apport d'histidine dans le milieu de croissance. Ces souches mutantes étaient stables et n'ont jamais présenté de réversion vers le type sauvage. Les deux mutants ont été croisés ensemble (*hist-1^a* X *hist-1^b*) et les milliers d'ascospores obtenues ont été étalées sur un milieu dépourvu d'histidine (milieu minimum). La plupart des ascospores n'ont pas poussé mais on a observé trois colonies. Utilisez un schéma pour expliquer l'origine de ces rares colonies prototrophes.
22. Nous sommes en 1868. Vous êtes un jeune et habile fabricant de lentilles optiques travaillant à Vienne. Avec les dernières lentilles que vous avez créées, vous venez juste de construire un microscope dont le pouvoir de résolution est bien supérieur à celui de tous les autres instruments contemporains disponibles. Au cours de vos tests de ce microscope, vous avez observé des cellules de testicules de sauterelles et vous avez été fasciné par le comportement d'étranges structures allongées que vous avez vues à l'intérieur des cellules en cours de division. Un jour, à la bibliothèque, vous lisez

Hidden page

27.



28. Chez la drosophile, les gènes déterminant la taille des ailes (phénotype : aile grande ou vestigiale) et la longueur des soies (phénotype : soie longue ou courte) sont transmis indépendamment. Les croisements suivants ont été réalisés :

Croisements parentaux	Descendants			
	Grande, longue	Grande, courte	Vestigiale, longue	Vestigiale, courte
Grande, longue × vestigiale, courte	246	0	250	0
Grande, longue × vestigiale, longue	302	98	297	101
Grande, longue × grande, courte	150	148	48	51
Grande, courte × vestigiale, longue	89	93	90	94

a. Déduisez-en les relations de dominance dans les phénotypes concernés.

b. Déduisez-en les génotypes des parents dans les quatre croisements.

29. Chez la drosophile, les gènes codant le corps ébène (*e*) et les soies épaisses (*s* pour *stubby* en anglais) sont liés sur le même bras du chromosome 2. Les mouches de génotype + *s* / *e* + se développent comme le type sauvage, mais présentent occasionnellement deux types différents d'anomalies inattendues sur leurs corps. La première anomalie est constituée par deux zones adjacentes, l'une portant des soies épaisses et l'autre de couleur ébène. La deuxième anomalie est constituée de taches isolées de couleur ébène.

a. Représentez par des schémas, les origines probables de ces deux types d'anomalies inattendues.

b. Expliquez pourquoi il n'y a pas de sites isolés de soies épaisses.

Hidden page

6

L'INTERACTION DES GÈNES



Les couleurs des poivrons sont déterminées par l'interaction de plusieurs gènes. Un allèle *Y* entraîne l'élimination précoce de la chlorophylle, ce qui n'est pas le cas de *y*. *R* détermine la couleur rouge et *r* les pigments caroténoïdes jaunes. Les allèles *c1* et *c2* de deux gènes différents régulent en aval les quantités de pigments caroténoïdes, donnant lieu à des couleurs moins vives. La couleur orange correspond à une régulation en aval du rouge. Le marron correspond à du vert additionné de rouge. Le jaune pâle est dû à une régulation en aval du jaune.

(Anthony Griffiths.)

POUR L'ESSENTIEL

- Lorsque deux allèles mutants récessifs distincts sont réunis dans une cellule et que le phénotype est mutant, les mutations appartiennent au même gène.
- Lorsque des allèles mutants récessifs sont réunis dans une cellule et que le phénotype est de type sauvage, les mutations sont situées dans des gènes différents.
- Les phénotypes de certains hétérozygotes révèlent des types de dominance autre que la dominance complète.
- Certains allèles mutants peuvent tuer l'organisme qui les abrite.
- La plupart des caractères sont déterminés par des ensembles de gènes qui interagissent les uns avec les autres et avec l'environnement.
- Les proportions modifiées de monohybrides (hétérozygotes pour un gène) révèlent des interactions alléliques.
- Les proportions modifiées de dihybrides (hétérozygotes pour deux gènes) révèlent des interactions entre gènes.

DES GÈNES AUX PHÉNOTYPES

Dans certains types d'expériences, les généticiens agissent comme si les gènes étaient isolés. À chaque fois que l'on réalise un croisement entre une souche à yeux rouges (type sauvage) de drosophile et une souche mutante à yeux blancs, chaque parent et chaque descendant contiennent quelque 20 000 gènes. Pourtant dans un seul croisement, nous ne remarquons que la particularité génétique qui différencie les mouches et qui nous intéresse — la couleur de l'œil de chaque mouche. En réalité, les gènes n'agissent pas de façon isolée. Les protéines et les ARN qu'ils codent participent à des suites de réactions spécifiques qui reçoivent également des informations des produits de nombreux autres gènes. De plus, l'expression d'un seul gène dépend d'un grand nombre de facteurs, y compris ceux constituant le fonds génétique des mouches et une vaste gamme de conditions environnementales — température, nourriture, densité de population, etc. (La dépendance de l'expression des gènes vis-à-vis des facteurs environnementaux a été traitée au chapitre 1, dans le paragraphe sur la norme de réaction).

L'action d'un gène est un terme qui recouvre une série très complexe d'événements. Il n'existe probablement aucun cas pour lequel nous comprenions tous les événements qui se produisent entre l'expression d'un gène et le phénotype correspondant dans un organisme. Dans ce chapitre, nous envisagerons certaines des méthodes utilisées par les généticiens pour démêler l'écheveau des mécanismes qui relient le gène au phénotype final.

Nous pouvons commencer par énoncer quelques généralités sur la complexité de l'action des gènes.

1. La relation entre les gènes et les phénotypes peut être plurivalente.

Cette relation s'appelle la *pléiotropie*. La pléiotropie se déduit du fait que lorsqu'on sélectionne des mutations pour leur effet sur un caractère spécifique, on s'aperçoit souvent qu'elles affectent d'autres caractères de l'organisme. Ceci laisse penser qu'il existe des voies physiologiques apparentées, qui contribuent à un phénotype similaire dans plusieurs tissus. Par exemple, la mutation blanche de la couleur de l'œil chez la drosophile résulte d'une déficience de pigmentation, non seulement dans les yeux composés, mais également dans les ocelles (yeux simples), les enveloppes de tissu entourant les gonades mâles et les tubules de Malpighi (les reins des mouches). Dans tous ces tissus, la formation de pigments nécessite l'absorption des précurseurs des pigments dans les cellules. L'allèle blanc entraîne un défaut de cette absorption, bloquant ainsi la formation de pigment dans tous ces tissus.

La pléiotropie implique souvent des événements multiples dont le lien physiologique n'est pas évident à première vue. Par exemple, la mutation dominante *Dichaete* (+ double -) chez la drosophile est responsable de la

disposition latérale et écartée des ailes mais enlève également des poils sur le dos de la mouche ; de plus, la mutation est létale lorsqu'elle est homozygote. Cet exemple montre une limite réelle à la façon dont on définit des mutations récessives ou dominantes. En réalité, une mutation peut à la fois être récessive et dominante, selon l'aspect du phénotype pléiotrope considéré. En général, la terminologie génétique n'est pas à même de définir ce niveau de pléiotropie et de complexité dans un symbole. C'est pourquoi, comme nous le verrons plus loin, il y a un certain aspect arbitraire ou historique dans la dénomination des allèles. Dans ce chapitre, nous nous intéresserons à certaines variations sur le thème général de la dominance et de la récessivité, ainsi qu'à certaines des nomenclatures traditionnelles utilisées pour désigner les allèles.

2. La relation entre les phénotypes et les gènes peut être plurivalente.

Cette idée, inverse de la précédente, est basée sur le fait que de nombreux gènes différents peuvent affecter un même phénotype. Cela est facile à comprendre pour un caractère tel que la couleur de l'œil, dans lequel interviennent des voies métaboliques complexes avec de nombreuses étapes enzymatiques, chacune codée par un ou plusieurs produits de gènes. Chez la drosophile, on estime que 100 gènes ou plus contribuent à la pigmentation de l'œil composé. Pour des processus plus complexes tels que le développement de la structure de l'œil composé, plusieurs centaines de gènes contribuent probablement au processus.

L'un des buts de l'analyse génétique est d'identifier tous les gènes qui affectent un phénotype spécifique et de comprendre leurs rôles génétique, cellulaire, développemental et moléculaire. Pour cela, nous devons disposer de moyens pour trier les mutations et les gènes. Dans le premier paragraphe de ce chapitre, nous allons envisager la façon dont nous pouvons utiliser l'analyse génétique pour déterminer si deux mutants résultent de mutations dans le même gène (c'est-à-dire s'il s'agit d'allèles), ou dans des gènes différents. Plus loin dans ce chapitre, nous verrons de quelle façon l'analyse génétique peut être utilisée pour faire des déductions sur les interactions des gènes dans les processus du développement et les voies biochimiques.

UN TEST DE DIAGNOSTIC POUR LES ALLÈLES

De nombreux programmes de recherche génétique sont en cours pour essayer de comprendre les gènes qui contribuent à un processus biologique particulier. Une telle analyse commence par l'isolement de phénotypes mutants apparentés, concernant un processus particulier. Par exemple, si un généticien s'intéresse aux gènes déterminant la

locomotion chez un ver nématode, la dissection génétique consiste d'abord à isoler un groupe de mutants différents présentant une locomotion déficiente. L'une des tâches importantes consiste à déterminer combien de gènes différents sont représentés par les allèles qui déterminent les phénotypes apparentés, car ceci définit l'ensemble des gènes qui affectent le processus étudié. Il est donc nécessaire de disposer d'un test pour découvrir si les mutations sont des allèles d'un même gène ou si elles touchent des gènes différents.

Le test d'allélisme qui a l'application la plus large est le **test de complémentation**, illustré dans l'exemple suivant. Considérons une espèce de campanule chez laquelle la couleur de la fleur de type sauvage est bleue. Supposons qu'après lui avoir fait subir des radiations mutagènes, nous ayons induit trois mutants à pétales blancs et qu'il s'agisse de souches homozygotes de lignée pure. Nous pouvons appeler ces souches mutantes \$, & et ¥, en utilisant des symboles de monnaies pour éviter d'influencer notre jugement en ce qui concerne la dominance. Lorsqu'on les croise avec le type sauvage, tous les mutants donnent les mêmes résultats dans la F_1 et la F_2 , comme indiqué ci-dessous :

Blanc \$ \times bleu $\rightarrow F_1$, bleues (toutes) $\rightarrow F_2$, $\frac{3}{4}$ bleues, $\frac{1}{4}$ blanches

Blanc & \times bleu $\rightarrow F_1$, bleues (toutes) $\rightarrow F_2$, $\frac{3}{4}$ bleues, $\frac{1}{4}$ blanches

Blanc ¥ \times bleu $\rightarrow F_1$, bleues (toutes) $\rightarrow F_2$, $\frac{3}{4}$ bleues, $\frac{1}{4}$ blanches

Dans chaque cas, les résultats montrent que l'état du mutant est déterminé par l'allèle récessif d'un seul gène. Cependant, s'agit-il de trois allèles d'un seul gène ou de deux ou encore de trois gènes ? On peut répondre à la question en se demandant si les mutants complètent les uns avec les autres. Définissons tout d'abord la notion de complémentation.

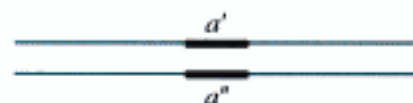
MESSAGE

La complémentation est la production d'un phénotype de type sauvage lorsque deux allèles mutants récessifs sont réunis dans la même cellule.

(La démonstration de la nature récessive de chaque mutant est un résultat essentiel qui nous permet de poursuivre l'étude par un test de complémentation. Les mutations dominantes ne peuvent être utilisées dans les tests de complémentation).

Dans le cas d'un organisme diploïde, le test de complémentation est réalisé en croisant deux à deux des mutants récessifs homozygotes et en regardant si la descendance est ou non de type sauvage. Si les mutations récessives représentent des allèles du même gène, alors à l'évidence, ces derniers ne complèteront pas car ils représentent tous deux la fonction perdue du gène. On peut désigner de tels allèles par a' et a'' , en utilisant les

primes pour faire la différence entre deux allèles mutants distincts d'un gène, dont l'allèle de type sauvage est a . Ces allèles pourraient avoir des sites mutants différents mais seraient fonctionnellement identiques. L'hétérozygote a'/a'' serait :



Cependant, deux mutations récessives dans des gènes différents auraient une fonction de type sauvage assurée par leurs allèles de type sauvage respectifs. Nous pouvons dans ce cas appeler les gènes $a1$ et $a2$, d'après leurs allèles mutants. Les hétérozygotes seraient $a1/+$; $+/a2$ (gènes non liés génétiquement) ou $a1 +/a2$ (gènes liés) et pourraient être schématisés de la façon suivante :

Gènes non liés



Gènes liés



Revenons à l'exemple de la campanule et croisons les mutants ensemble pour un test de complémentation. Supposons que les résultats des croisements des mutants \$, & et ¥ soient les suivants :

Blanche \$ \times blanche & $\rightarrow F_1$, (toutes blanches)

Blanche \$ \times blanche ¥ $\rightarrow F_1$, (toutes bleues)

Blanche & \times blanche ¥ $\rightarrow F_1$, (toutes bleues)

D'après ces résultats, nous pouvons conclure que les mutants \$ et & doivent être dus aux allèles d'un gène (disons $w1$) car ils ne complètent pas. ¥, au contraire, doit résulter d'un allèle mutant d'un autre gène ($w2$).

L'explication moléculaire de tels résultats se trouve souvent dans les voies biochimiques de la cellule. Comment la complémentation fonctionne-t-elle au niveau moléculaire ? Bien que par convention on dise que des mutants complètent, en fait, les agents actifs dans la complémentation sont les protéines produites par les allèles de type sauvage. La couleur bleue normale de la fleur est due à un pigment bleu appelé *anthocyanine*. Les

Hidden page

Une mutation dans l'un ou l'autre des gènes à l'état homozygote conduira à l'accumulation d'un précurseur, qui rendra simplement la fleur blanche. On peut à présent écrire les génotypes de la façon suivante :

$$\begin{array}{ll} \$ & w1_s/w1_s, w2^+/w2^+ \\ \& & w1_k/w1_k, w2^+/w2^+ \\ \gamma & w1^+/w1^+, w2_\gamma/w2_\gamma \end{array}$$

Cependant en pratique les symboles placés en indice seraient supprimés et les génotypes écrits ainsi :

$$\begin{array}{ll} \$ & w1/w1, w2^+/w2^+ \\ \& & w1/w1, w2^+/w2^+ \\ \gamma & w1^+/w1^+, w2/w2 \end{array}$$

Une F_1 provenant de $\$ \times \&$ sera donc :

$$w1/w1, w2^+/w2^+$$

et aura deux allèles déficitants pour *w1*. Elle sera donc bloquée à l'étape 1. Même si l'enzyme de l'étape 2 est parfaitement fonctionnelle, elle n'aura pas de substrat sur lequel agir. Aucun pigment bleu ne sera donc produit et le phénotype sera blanc.

Toutefois, la F_1 provenant des autres croisements possèdera les allèles de type sauvage des deux enzymes, nécessaires pour assurer les interconversions jusqu'au

produit bleu final. Leurs génotypes seront :

$$w1^+/w1, w2^+/w2$$

Nous voyons donc ici en quoi la complémentation est effectivement un résultat de l'interaction coopérative des allèles de *type sauvage* des deux gènes. La figure 6-1 présente un résumé schématique de l'interaction de mutants blancs qui complémentent et qui ne complémentent pas.

Dans le cas d'un organisme haploïde, le test de complémentation ne peut être réalisé par croisement. Chez les champignons, il existe un test alternatif qui consiste à fabriquer un **hétérocaryon** (figure 6-2). Les cellules de champignons fusionnent facilement et lorsque deux souches différentes fusionnent, des noyaux des deux souches peuvent occuper la même cellule, que l'on appelle alors un *hétérocaryon* (du grec : différents noyaux). Les noyaux d'un hétérocaryon ne fusionnent généralement pas. En un sens, cela reproduit une sorte d'état diploïde. Supposons que dans des souches différentes nous ayons des mutations dans deux gènes distincts, conférant le même phénotype mutant, par exemple, un besoin en arginine. Nous pouvons appeler ces gènes *arg-1* et *arg-2*. Les deux souches, dont les génotypes peuvent être représentés par *arg-1*, *arg-2*⁺ et *arg-1*⁺, *arg-2* peuvent fusionner pour former un hétérocaryon chez lequel les deux noyaux sont dans le même cytoplasme :

Noyau 1: *arg-1* . *arg-2*⁺

Noyau 2: *arg-1*⁺ . *arg-2*

L'expression des gènes ayant lieu dans un cytoplasme commun (car les protéines sont synthétisées dans le cytoplasme), les deux allèles de type sauvage peuvent exercer

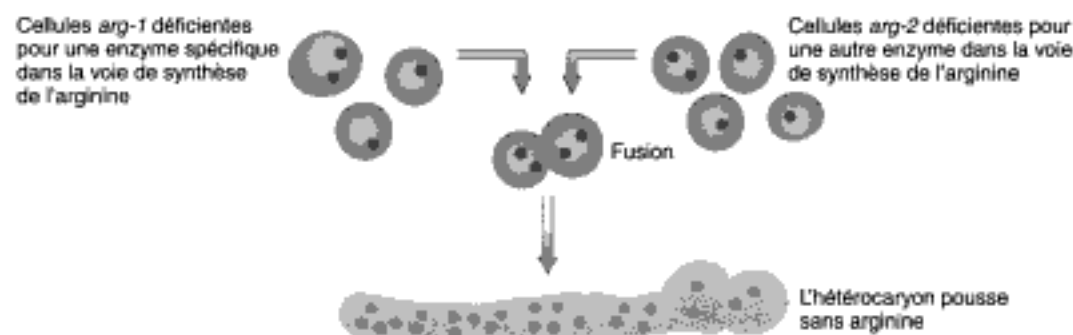


Figure 6-2 La formation d'un hétérocaryon chez *Neurospora*, démontrant à la fois la complémentation et la récessivité. Les cellules végétatives de ce champignon normalement haploïde peuvent fusionner, permettant aux noyaux des deux souches de se mélanger dans le même cytoplasme. Si chaque souche est bloquée en un endroit différent d'une même voie métabolique, comme c'est le cas pour les mutants *arg-1* et *arg-2*, toutes les fonctions seront assurées dans l'hétérocaryon et le champignon *Neurospora* poussera ; en d'autres termes, il y aura complémentation. La croissance est possible grâce à la présence d'un gène de type sauvage. Les gènes mutants *arg-1* et *arg-2* doivent donc être récessifs.

leur effet dominant et coopérer pour produire un hétérozygote de phénotype sauvage. En d'autres termes, les deux mutations complètent exactement de la même manière que s'il s'agissait d'un diploïde. Si les mutations correspondaient aux allèles d'un même gène, il n'y aurait pas eu de complément.

MESSAGE

Lorsque deux allèles mutants récessifs qui proviennent de deux origines distinctes et qui produisent des phénotypes récessifs similaires ne complètent pas, les allèles doivent correspondre au même gène. S'ils complètent, ils appartiennent à des gènes différents.

LES INTERACTIONS ENTRE LES ALLÈLES D'UN MÊME GÈNE

Les allèles d'un même gène peuvent interagir de plusieurs façons différentes au niveau fonctionnel, produisant des variations dans le type de dominance et des effets parfois bien différents sur le phénotype, selon les combinaisons alléliques.

La dominance incomplète

Les belles-de-nuit sont des plantes originaires d'Amérique tropicale. Leur nom vient du fait qu'elles s'ouvrent vers la fin de l'après-midi. Lorsqu'une belle-de-nuit de type sauvage avec des pétales rouges est croisée avec une lignée pure dont les pétales sont blancs, alors la F_1 a des pétales roses. Si une F_1 est produite en croisant les plantes de la F_1 ensemble, alors le résultat est :

- $\frac{1}{4}$ des plantes ont des pétales rouges
- $\frac{1}{2}$ des plantes ont des pétales roses
- $\frac{1}{4}$ des plantes ont des pétales blancs

Le rapport étant de 1 : 2 : 1 dans la F_2 , nous pouvons déduire qu'il s'agit de la transmission de deux allèles d'un même gène. Cependant les hétérozygotes (la F_1 et la moitié de la F_2) ont des phénotypes intermédiaires, ce qui suggère un type incomplet de dominance. En inventant des symboles pour les allèles, nous pouvons dresser ainsi la liste des génotypes des belles-de-nuit de cette expérience : c^+/c^+ (rouge), c/c (blanc) et c^+/c (rose). La **dominance incomplète** décrit la situation dans laquelle le phénotype d'un hétérozygote est intermédiaire entre ceux des deux homozygotes sur une échelle de mesure quantitative. La figure 6-3 indique les termes correspondant à

toutes les positions théoriques sur l'échelle, mais en pratique, il est difficile de déterminer exactement l'endroit où l'hétérozygote est situé sur l'échelle. Au niveau moléculaire, la dominance incomplète est généralement provoquée par un effet quantitatif du nombre de « doses » d'un allèle de type sauvage ; deux doses produisent davantage de transcrits fonctionnels et donc de protéines fonctionnelles ; une dose produit moins de transcrits et de produits et aucune dose ne débouche sur aucun transcrit ni protéine fonctionnels. Dans le cas de la dominance complète, chez l'hétérozygote sauvage/mutant, soit la moitié de la quantité normale du transcrit et du produit permettent d'assurer une fonction cellulaire normale (le gène est haplo-suffisant), soit l'allèle de type sauvage est « régulé en amont » pour augmenter la concentration du transcrit jusqu'à un niveau normal.

La codominance

Les groupes sanguins humains ABO sont déterminés par trois allèles d'un gène, qui présentent plusieurs types d'interaction aboutissant aux quatre types sanguins du système ABO. Les séries alléliques comprennent trois allèles principaux, I^A , I^B et I^O , mais il est évident que chaque personne ne peut posséder que deux des trois allèles (ou deux copies d'un même allèle). Il y a six génotypes différents, trois homozygotes et trois types distincts d'hétérozygotes :

Génotype	Type sanguin
I^A/I^A , I^A/i	A
I^B/I^B , I^B/i	B
I^A/I^B	AB
i/i	O

Dans cette série allélique, les allèles I^A et I^B déterminent chacun un antigène unique, déposé à la surface des globules rouges. Ce sont deux formes de la même

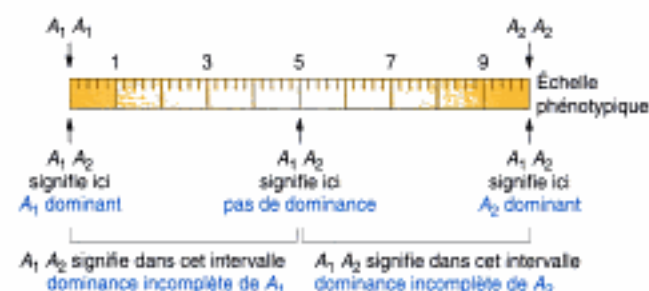


Figure 6-3 Un résumé des relations de dominance. La règle représente une sorte de mesure phénotypique, telle que la quantité de pigment.

protéine. En revanche, l'allèle i ne produit aucune protéine antigénique de ce type. Dans les génotypes P/i et P^+/i , les allèles P et P^+ sont complètement dominants sur i . Cependant, dans le génotype P/P^+ , chacun des allèles produit son propre antigène. On dit donc qu'ils sont **codominants**.

L'anémie à cellules falciformes, une maladie humaine, donne un aperçu intéressant de la dominance. Le gène concerné affecte la molécule d'hémoglobine qui transporte l'oxygène et est le principal constituant des globules rouges. Les trois génotypes ont des phénotypes différents comme on le voit ci-dessous :

Hb^A/Hb^A	Normal ; les globules rouges ne prennent jamais la forme de faucille.
Hb^S/Hb^S	Anémie grave, souvent mortelle ; l'hémoglobine anormale provoque la déformation des globules rouges en faucille (= en forme de faucille).
Hb^A/Hb^S	Pas d'anémie ; les globules rouges n'adoptent la forme de faucille que lorsque la concentration en oxygène est faible.

La figure 6-4 présente des cellules en forme de faucille. Si l'on considère la présence ou l'absence d'anémie, l'allèle Hb^A est à l'évidence dominant. Par rapport à la forme des globules rouges cependant, il y a dominance incomplète. Pour finir, comme nous le verrons désormais, par rapport à l'hémoglobine elle-même, il y a codominance. Les allèles Hb^A et Hb^S codent en réalité deux formes distinctes d'hémoglobine qui diffèrent par un seul acide aminé ; ces deux formes sont synthétisées chez l'hétérozygote. Les différentes formes d'hémoglobine peuvent être visualisées par électrophorèse, une technique qui sépare les macromolécules en fonction de leur charge ou de leur taille (figure 6-5). Il s'avère que les formes A et S de l'hémoglobine ont des charges différentes, qui permettent de les séparer par électrophorèse (figure 6-6). Nous voyons que les personnes homozygotes normales possèdent un type d'hémoglobine (A) et que les personnes anémiques ont un type S qui se dé-

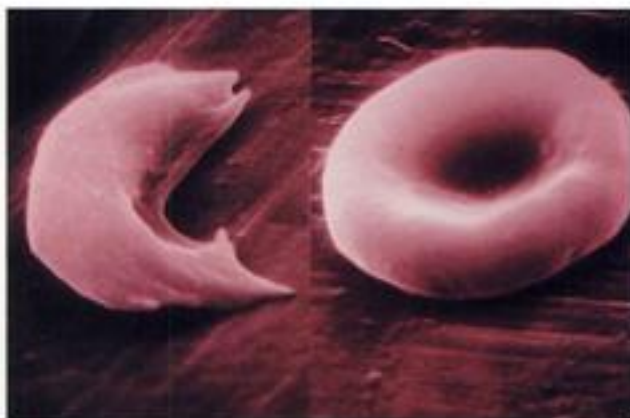


Figure 6-4 Micrographie électronique de globules rouges issus d'un individu atteint d'anémie à cellules falciformes. Quelques cellules arrondies (à droite) semblent quasiment normales. (© Stan Flegler/ Visuals Unlimited).

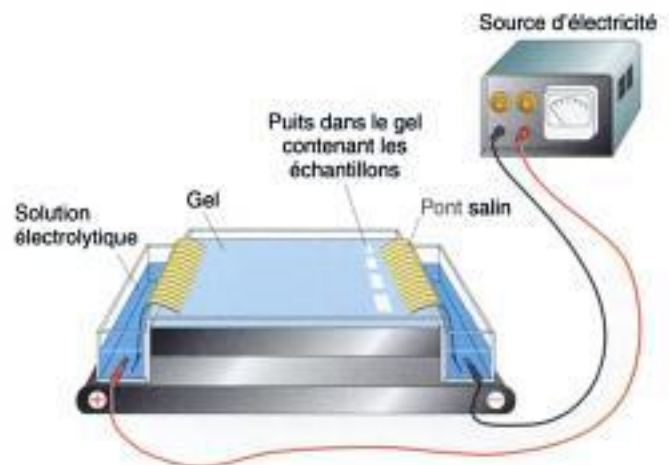


Figure 6-5 Appareil servant à l'électrophorèse. Chaque échantillon est placé dans un puits à l'intérieur d'une plaque gélatineuse (un gel). Les molécules des échantillons migrent à des distances différentes dans le gel, selon leur charge électrique. Plusieurs échantillons sont testés en même temps (un dans chaque puits). Les positions auxquelles les molécules ont migré sont ensuite révélées par coloration.

place plus lentement dans le champ électrique. Les hétérozygotes possèdent les deux types, A et S. En d'autres termes, il y a codominance au niveau moléculaire.

L'anémie à cellules falciformes illustre le fait que les termes *dominance*, *dominance incomplète* et *codominance* sont quelque peu arbitraires. Le type de dominance que l'on en déduit dépend du niveau phénotypique auquel on effectue les observations — moléculaire, cellulaire ou de l'organisme. Il faut porter la même attention à un grand nombre des catégories utilisées par les scientifiques pour classer les structures et les réactions. Ces catégories sont imaginées par l'homme pour faciliter les analyses.

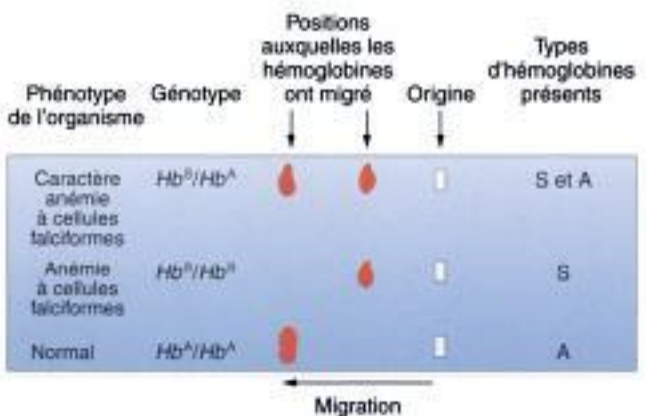


Figure 6-6 Électrophorèse de l'hémoglobine provenant d'un individu atteint d'anémie à cellules falciformes, un hétérozygote (appelé *caractère anémie à cellules falciformes*) et un individu normal. Les taches montrent les positions auxquelles les hémoglobines migrent dans le gel.

MESSAGE

Le type de dominance est déterminé par les fonctions moléculaires des allèles d'un gène et par le niveau d'analyse.

Le trèfle est le nom courant des plantes du genre *Trifolium*. Il en existe de nombreuses espèces. Certaines sont originaires d'Amérique du Nord, d'autres sont apparues comme des mauvaises herbes. Beaucoup de recherches génétiques ont été effectuées avec du trèfle blanc, qui présente une variation considérable parmi les individus affectés du motif curieux en V (ou chevron) sur les feuilles. La figure 6-7 montre que les différentes formes de chevron (et l'absence de chevron) sont déter-



Figure 6-7 Des allèles multiples déterminent le motif en chevron sur les feuilles de trèfle blanc. Le génotype des plantes est indiqué sous chacune d'elles. (Adapté d'une photographie de W. Ellis Davies).

minées par des allèles multiples. Dans cet exemple, nous avons affaire à un polymorphisme génétique. C'est pourquoi le symbolisme allèle mutant/sauvage n'est pas utilisé. Étudiez la figure pour déterminer le type de dominance de chaque allèle dans les différentes combinaisons. Faites la liste des allèles de façon à mettre en évidence leurs relations de dominance les uns avec les autres. Y a-t-il des incertitudes ? L'observation nous permet-elle de dire quelque chose sur la dominance ou la récessivité de l'allèle *v* ?

Les allèles létaux

Les souris normales de type sauvage ont un pelage avec une pigmentation assez sombre. Une mutation appelée *jaune* (*yellow* en anglais, une couleur de pelage plus claire) illustre une interaction allélique intéressante. Si une souris jaune est croisée avec une souris homozygote de type sauvage, on observe toujours un rapport 1 : 1 entre les souris jaunes et les souris de type sauvage de la descendance. Cette observation suggère (1) qu'un seul gène avec deux allèles détermine ces phénotypes alternatifs, (2) que la souris jaune était hétérozygote pour ces allèles et (3) que l'allèle responsable du pelage jaune est dominant sur l'allèle de la couleur normale. Toutefois, si deux souris jaunes sont croisées l'une avec l'autre, le résultat est toujours le suivant :

$$\text{Jaune} \times \text{jaune} \longrightarrow \frac{2}{3} \text{ jaune}, \frac{1}{3} \text{ de type sauvage}$$

Deux éléments intéressants sont à noter dans ces résultats. Premièrement, le rapport phénotypique 2 : 1 est différent de celui attendu pour l'autofécondation d'un monohybride. Deuxièmement, puisqu'aucun croisement jaune x jaune ne produit jamais que des descendants jaunes comme ce serait le cas si chaque parent était homozygote, il semble qu'il soit impossible d'obtenir des souris jaunes homozygotes.

L'explication d'un tel résultat est que toutes les souris jaunes sont hétérozygotes pour un allèle particulier. On s'attendrait à ce qu'un croisement entre deux hétérozygotes conduise à un rapport génotypique monohybride 1 : 2 : 1. Toutefois, si toutes les souris de l'une de ces classes homozygotes mouraient avant la naissance, les souris nées présenteraient alors un rapport 2 : 1 entre les hétérozygotes et les homozygotes survivantes. L'allèle A^1 pour le pelage jaune est dominant sur l'allèle de type sauvage A en ce qui concerne son effet sur la couleur, mais A^1 agit comme un allèle récessif **léta** par rapport au caractère que nous appelons la *viabilité*. Par conséquent, une souris de génotype homozygote A^1/A^1 meurt donc avant la naissance et n'est pas observée parmi les descendants. Toutes les souris jaunes survivantes doivent donc être hétérozygotes A^1/A . Un croisement entre les souris jaunes conduira donc toujours aux résultats suivants :

Hidden page

détermine une protéine anormale du cartilage et l'effet de cette anomalie est exprimé phénotypiquement dans plusieurs organes différents, entraînant des symptômes létaux, comme le montre la figure 6-10. L'anémie à cellules falciformes, dont nous avons parlé plus haut, en est un autre exemple.

Qu'un allèle soit léthal ou non dépend souvent de l'environnement dans lequel l'organisme se développe. Alors que certains allèles sont létaux dans la plupart des environnements, d'autres sont viables dans un environnement, mais létaux dans un autre. Les maladies héréditaires humaines en sont des exemples. La mucoviscidose est une maladie qui serait létale sans traitement et les individus atteints de PCU ne survivaient certainement pas dans un environnement naturel, dans lequel un régime spécial serait impossible. D'autre part, la majorité des allèles favorisés et sélectionnés par les éleveurs d'animaux et les cultivateurs seraient presque quasiment éliminés dans la nature, à la suite de la compétition entre membres de populations naturelles. Les variétés modernes de graines sont de bons exemples ; seule une culture attentive par le fermier a maintenu de tels allèles dans notre intérêt.

Les généticiens rencontrent couramment des situations dans lesquelles les rapports phénotypiques attendus

sont régulièrement décalés dans un sens, par une viabilité réduite à cause d'un allèle. Par exemple dans le croisement $A/a \times a/a$, nous prévoyons un rapport phénotypique dans la descendance de 50 pour cent de A/a et de 50 pour cent de a/a , mais nous pouvons observer des rapports tels que 55 % : 45 % ou 60 % : 40 %. Dans ce cas, le phénotype récessif est dit *subvital* ou *semi-létal*, car la létalité est exprimée seulement chez quelques individus. La létalité peut donc aller de 0 à 100 pour cent, selon le gène lui-même, le reste du génome et l'environnement.

MESSAGE

Un gène peut avoir plusieurs états ou formes — c'est ce qu'on appelle des allèles multiples. On dit que les allèles constituent une série allélique. Les membres d'une même série peuvent présenter différents degrés de dominance les uns par rapport aux autres.

L'INTERACTION DES GÈNES CONDUIT À DES PROPORTIONS MODIFIÉES DE DIHYBRIDES

L'analyse génétique permet d'identifier les gènes qui interagissent pour spécifier une propriété biologique donnée. On sait que deux gènes qui interagissent produisent des *proportions modifiées de dihybrides*. Il existe de nombreuses sortes d'interaction, ce qui conduit à des modifications très variées. Il est important de distinguer les gènes qui interagissent dans des voies biologiques distinctes, de ceux qui appartiennent à la même voie.

Les gènes en interaction dans des voies biologiques distinctes

Un exemple simple mais frappant de l'interaction des gènes est la transmission de la couleur de la peau chez les couleuvres *Elaphe guttata*. La couleur normale est un motif répété de camouflage, constitué d'une succession de noir et d'orange, comme le montre la figure 6-11a. Le phénotype est produit par deux pigments séparés, tous deux sous contrôle génétique. Un gène détermine le pigment orange et les allèles que nous allons voir seront o^+ (présence de pigment orange) et o (absence de pigment orange). Un autre gène détermine le pigment noir, avec les allèles b^+ (présence de pigment noir (*black* en anglais)) et b (absence de pigment noir). Ces deux gènes ne sont pas liés génétiquement. Le motif naturel est produit par le génotype $o^+/- ; b^+/-$. Un serpent $o/o ; b^+/-$ est noir car il n'a pas de pigment orange (figure 6-11b). Un serpent $o^+/- ; b/b$ est orange car il n'a pas de pigment noir (figure 6-11c). Le double homozygote récessif $o/o ; b/b$ est albinos,

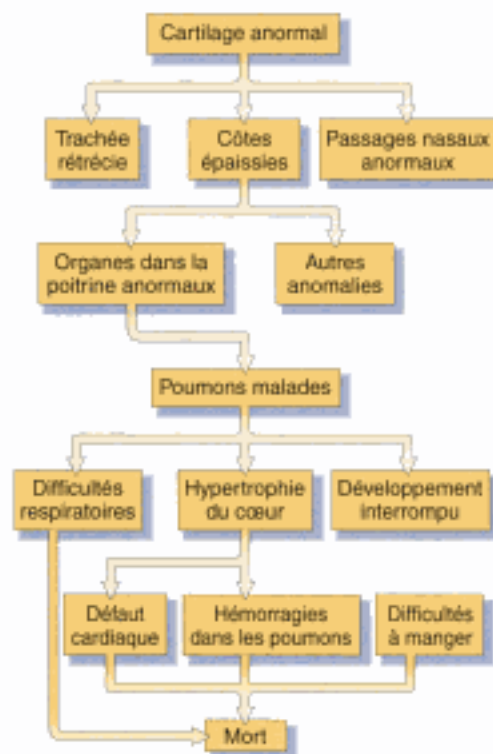
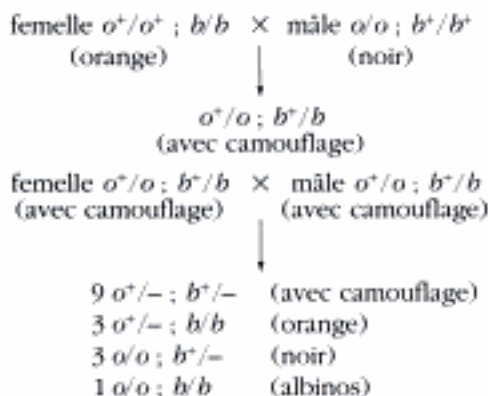


Figure 6-10 Schéma montrant de quelle façon un allèle léthal particulier entraîne la mort chez les rats. (D'après I. M. Lerner et W. J. Libby, *Heredity, Evolution and Society*, 2^e édition. W. H. Freeman and Company, 1976 ; d'après H. Gruneberg).

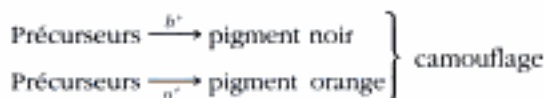


Figure 6-11 L'analyse des gènes de la pigmentation de la peau chez la couleuvre *Elaphe guttata*. Le type sauvage (a) a un motif de pigmentation de la peau, composé de pigments noirs et orange. Le gène o^+ détermine une enzyme dans la voie de synthèse du pigment orange ; lorsque cette enzyme est déficiente (o/o), aucun pigment orange n'est fabriqué et le serpent est noir (b). Un autre gène, b^+ , détermine une enzyme dans la voie de synthèse du pigment noir (*black* en anglais) ; lorsque cette enzyme est déficiente (b/b), le serpent est orange (c). Lorsque les deux enzymes sont déficientes, le serpent est albinos (d). Les quatre génotypes homozygotes sont donc (a) $o^+/o^+ ; b^+/b^+$, (b) $o/o ; b^+/b^+$, (c) $o^+/o^+ ; b/b$ et (d) $o/o ; b/b$. Un croisement de (a) x (d) ou de (b) x (c) donnerait une F_1 dihybride de type sauvage et un rapport 9 : 3 : 3 : 1 des quatre phénotypes dans la F_2 . (Anthony Griffiths).

comme le montre la figure 6-11d. Remarquez cependant que la couleur rose pâle de l'albinos est due à un autre pigment, l'hémoglobine du sang, qui est visible au travers de la peau de ce serpent lorsque les autres pigments sont absents. De plus, le serpent albinos montre clairement qu'un autre élément se superpose à la pigmentation de la peau — le motif répété à l'intérieur et autour duquel les pigments sont déposés. Puisqu'il y a deux gènes dans ce système, nous obtenons un mode de transmission typique d'un dihybride et les quatre phénotypes forment un rapport 9 : 3 : 3 : 1 dans la F_2 . On pourrait réaliser une analyse type qui serait la suivante :



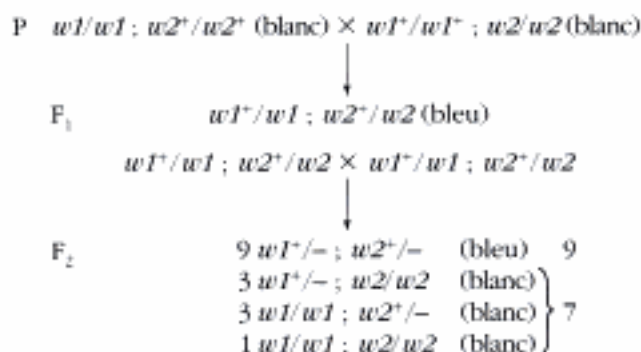
En résumé, on observe le rapport 9 : 3 : 3 : 1 de dihybrides parce que les mutations se trouvent dans deux voies biochimiques parallèles :



Lorsque les mutations appartiennent à la même voie biologique, on observe des rapports différents, comme nous allons le voir dans les exemples suivants.

Les gènes en interaction dans la même voie biologique

Nous avons simplement besoin ici de revenir à l'exemple de la couleur des pétales chez les campanules, que nous avons traité lors de la complémentation. La voie de synthèse se terminait par un pigment bleu et les intermédiaires étaient tous incolores. Deux lignées homozygotes différentes de campanules ont été croisées et la F_1 avait des fleurs bleues, montrant ainsi une complémentation. A quoi ressembleront les plantes de la F_2 issues du croisement des plantes de la F_1 ? La F_2 présente à la fois des plantes à fleurs bleues et des plantes à fleurs blanches (*white* en anglais) dans un rapport 9 : 7. Comment peut-on expliquer ces résultats ? Le rapport 9 : 7 est à l'évidence une modification du rapport 9 : 3 : 3 : 1 de dihybrides, la combinaison de 3 : 3 : 1 correspondant au 7. Le croisement des deux lignées blanches et les générations suivantes peuvent être représentés ainsi :

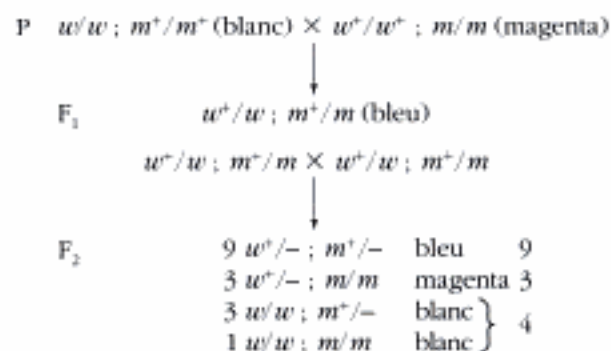


Les résultats montrent qu'une homozygotie pour l'allèle mutant récessif de l'un ou l'autre des gènes aboutit à une plante avec des pétales blancs. Pour manifester le phénotype bleu, une plante doit avoir au moins un allèle dominant des deux gènes.

Si un ou plusieurs intermédiaires de la voie de synthèse sont colorés, alors un rapport différent est produit dans la F_2 . Dans l'exemple suivant, concernant la plante *Collinsia parviflora*, la voie est la suivante :



Les gènes w et m ne sont pas liés génétiquement. Si des plantes homozygotes blanches et magenta sont croisées, les F_1 et F_2 seront ainsi :



On obtient finalement un rapport phénotypique de 9 : 3 : 4. Ce type d'interaction est appelé **épistasie**, qui signifie littéralement « être au-dessus ». En d'autres termes, un allèle d'un gène masque l'expression des allèles d'un autre gène. Dans cet exemple, l'allèle w est épistatique sur m^+ et m . À l'inverse, m^+ et m ne peuvent être exprimés qu'en présence de w^+ . Ici, un allèle récessif étant épistatique, il s'agit d'un cas d'**épistasie récessive**.

MESSAGE

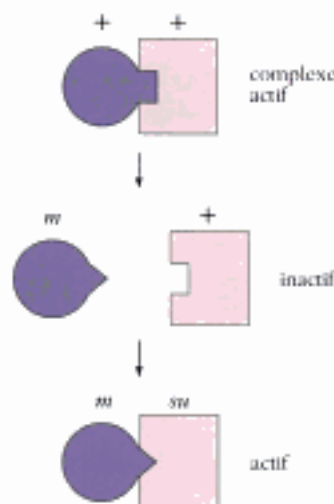
On déduit une épistasie lorsqu'un allèle d'un gène élimine l'expression des allèles d'un autre gène et exprime son propre phénotype à la place.

Hidden page

Le rapport global dans la F_2 est de 13 rouge : 3 violet. Ce rapport est caractéristique d'un suppresseur récessif qui agit sur une mutation récessive. Il existe des suppresseurs dominants et des suppresseurs récessifs.

On confond parfois la suppression avec l'épistasie. Cependant, puisqu'un suppresseur efface l'expression d'un allèle mutant et restaure le phénotype de type sauvage correspondant, le rapport modifié des dihybrides de la F_2 peut impliquer seulement deux phénotypes (normal et anormal) alors que dans le cas de l'épistasie, l'allèle épistatique introduit un troisième phénotype dans le rapport.

Comment les suppresseurs fonctionnent-ils au niveau moléculaire ? Il y a de nombreux mécanismes possibles. Un type très recherché est celui des **suppresseurs de non-sens**, qui agissent sur les mutations causées par l'apparition de codons stop (- non-sens -), au milieu de la séquence codant une protéine. Les mutants non-sens présentent une terminaison anticipée de la chaîne d'acides aminés. Toutefois, une mutation dans un anticodon d'ARNt qui permet à l'ARNt d'insérer un acide aminé au niveau d'un codon non-sens supprimera l'effet de la mutation en permettant à la synthèse protéique de se poursuivre au-delà du site de la mutation dans l'ARNm. Puisqu'il y a souvent répétition des gènes d'ARNt, ce suppresseur sera parfaitement viable. Un autre type de suppression est possible dans les cas d'interactions protéine-protéine. Si en temps normal, deux protéines s'emboîtent pour assurer une fonction cellulaire donnée, alors, lorsqu'une mutation entraîne le changement de forme d'une protéine, les deux protéines ne peuvent plus s'associer et la fonction disparaît. Cependant, un changement de forme compensatoire dans la seconde protéine peut agir comme un suppresseur pour restaurer l'association normale.



Enfin, dans des situations où une mutation provoque un blocage dans une voie métabolique, le suppresseur trouve le moyen de contourner le blocage, par exemple en acheminant des substances situées au-delà du blocage à partir de voies métaboliques apparentées.

Puisque l'interaction entre un suppresseur et son gène cible peut être prouvée, les généticiens recherchent délibérément des suppresseurs afin de reconstituer des groupes de gènes en interaction, qui affectent un processus ou une structure biologique. L'approche est relativement simple — réaliser à une grande échelle une mutagenèse à partir d'une lignée mutante (disons de génotype m) et rechercher simplement les rares individus de type sauvage. La plupart d'entre eux correspondront à des réversions (mutations inverses) m^+ , mais certains correspondront à une suppression ($m. su$) et pourront être distingués des précédents par les proportions des dihybrides, obtenues lors des croisements. Cette procédure peut très facilement être appliquée aux organismes haploïdes. Par exemple, si un grand nombre de cellules d'un mutant ayant besoin d'arginine (arg) sont étalées sur une boîte contenant un milieu de croissance dépourvu d'arginine, la plupart des cellules ne pousseront pas. En revanche, les cellules qui contiennent des réversions qui restaurent le type sauvage (arg^+) et des mutations suppressives ($arg. su$) pousseront et annonceront leur présence en formant des colonies visibles. Les colonies porteuses de la mutation suppressive peuvent être distinguées par un croisement avec le type sauvage, car des descendants ayant besoin d'arginine seront produits :

$$\begin{array}{rcl}
 arg. su \times arg^+. su^+ & & \\
 \downarrow & & \\
 \left. \begin{array}{l} arg^+. su^+ \\ arg^+. su \\ arg. su \end{array} \right\} & \text{n'ont pas besoin d'arginine} & \\
 arg. su^+ & \text{ont besoin d'arginine} &
 \end{array}$$

MESSAGE

Les suppresseurs effacent l'expression d'un allèle mutant d'un autre gène, restaurant le phénotype sauvage normal.

La couleur du pelage chez un modèle de mammifère, la souris

L'analyse de la couleur du pelage chez les mammifères est un splendide exemple de la façon dont des gènes différents coopèrent pour déterminer l'aspect global d'un pelage. La souris est un mammifère qui convient bien aux études génétiques car elle est petite et donc facile à entretenir en laboratoire et parce que son cycle de reproduction est court. C'est le mammifère le mieux étudié en ce qui concerne le déterminisme génétique de la couleur du pelage. Celui des autres mammifères est très proche de celui des souris, et pour cette raison, la souris sert de système modèle. Nous regarderons des exemples d'autres mammifères plus loin dans notre discussion. Cinq gènes

Hidden page

plupart des cas, le gène code une enzyme, la tyrosinase, qui produit la mélanine. Chez les lapins, un allèle de ce gène, l'allèle c^h (himalayen), impose un dépôt de pigments uniquement au niveau des extrémités du corps. Chez les souris, cet allèle produit également le phénotype appelé *Himalayen* et chez les chats, le phénotype appelé *Siamois* (figure 6-15). L'allèle c^h peut être considéré comme une version de l'allèle c avec une expression sensible à la chaleur. C'est seulement au niveau des extrémités les plus froides du corps que c^h est fonctionnel et peut conduire à la production de pigment. Dans les parties chaudes du corps, il est exprimé exactement comme l'allèle albinos c . Cet allèle montre clairement de quelle façon l'expression d'un allèle dépend de l'environnement.

Le gène D. Le gène D contrôle l'intensité du pigment spécifié par les autres gènes de la couleur du pelage. Les génotypes D/D et D/d permettent l'expression intégrale de la couleur chez les souris, mais d/d atténue la couleur, la rendant blanchâtre. L'effet est dû à une distribution inégale du pigment dans les poils. Toutes les couleurs de pelage, agouti atténué, cannelle atténuée, marron atténué et noir atténué sont possibles. Un gène avec un tel effet

est appelé **gène modificateur**. Chez les chevaux, l'allèle D présente une dominance incomplète. La figure 6-16 montre de quelle façon l'atténuation affecte l'aspect de chevaux alezan et bai. On observe fréquemment des cas d'atténuation de la couleur du pelage chez les chats domestiques.

L'action du gène modificateur peut être basée sur de nombreux mécanismes moléculaires différents. Dans un cas, elle implique des gènes régulateurs qui se fixent à la région en amont du gène, à côté du promoteur, et affectent le niveau de transcription. Les régulateurs positifs augmentent les taux de transcription et les régulateurs négatifs les abaissent. Considérons l'exemple de la régulation d'un gène G . G est l'allèle normal codant une protéine active, alors que g est un allèle porteur d'une mutation complète (« allèle nul », dû à une substitution d'une paire de bases) qui code une protéine inactive. Au niveau d'un locus non lié, R code une protéine régulatrice qui augmente le taux de transcription au niveau du locus G , alors que r conduit à une protéine qui permet seulement un niveau de transcription basal. Si un dihybride $G/g ; R/r$ subit une autofécondation, on obtient un rapport



(a)

Figure 6-15 Les allèles du gène C thermo-sensibles produisent des phénotypes similaires chez plusieurs mammifères différents. Ces allèles conduisent à une production très réduite ou nulle du pigment noir de mélanine, dans la peau recouvrant les parties les plus chaudes du corps. À des températures plus basses, telles que celles observées au niveau des extrémités du corps, la mélanine est synthétisée, produisant museau, oreilles, queue et bout des pattes plus sombres. (a) Souris himalayenne, (b) Chat siamois. Par ailleurs, les lapins himalayens (non représentés ici) sont souvent vendus comme animaux de compagnie. Tous les trois sont de génotype c^h/c^h . (Partie a, Anthony Griffiths ; partie b, Walter Chandoha).



(b)

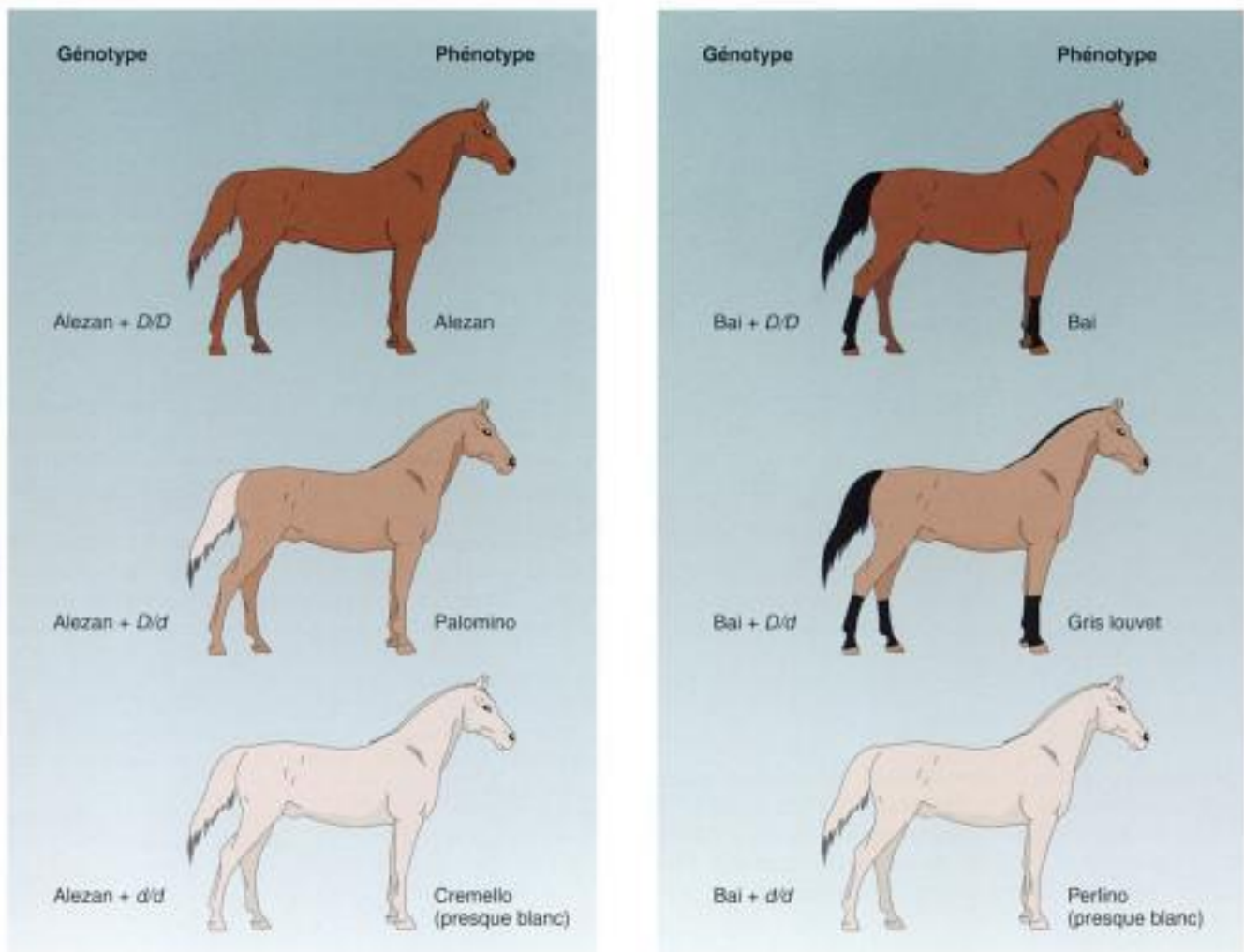


Figure 6-16 L'effet modificateur de l'allèle d'atténuation sur des génotypes alezan et bai chez les chevaux. Remarquez la dominance incomplète présentée par *D*. (D'après J. W. Ewans et coll., *The horse*, W. H. Freeman and Company, 1977).

de 9 : 3 : 4 pour les niveaux d'activité de la protéine, comme on le voit ci-dessous :

		Transcription	Activité protéique
9	$G^+ - ; R^+ -$	Élevée	Élevée
3	$G^+ - ; r/r$	Basse	Basse
3	$g/g ; R^+ -$	Élevée	Aucune
1	$g/g ; r/r$	Basse	Aucune

Le gène S. Le gène *S* détermine la présence ou l'absence de taches en contrôlant la migration de groupes de méla-

nocytes (cellules produisant des pigments) à la surface de l'embryon en cours de développement. Le génotype S^- ne donne aucun tache, tandis que s/s donne une apparence tachetée, appelée *pie* chez les souris et les chevaux. Cet aspect peut se superposer à n'importe laquelle des couleurs de pelage dont nous avons parlé précédemment, à l'exception des albinos bien entendu. On connaît également des mutations *pie* chez l'homme.

Nous voyons que l'aspect normal d'un pelage chez des souris sauvages est produit par un ensemble complexe de gènes en interaction déterminant le type de pigment, leur distribution dans chacun des poils et sur le corps de l'animal et la présence ou l'absence de pigments.

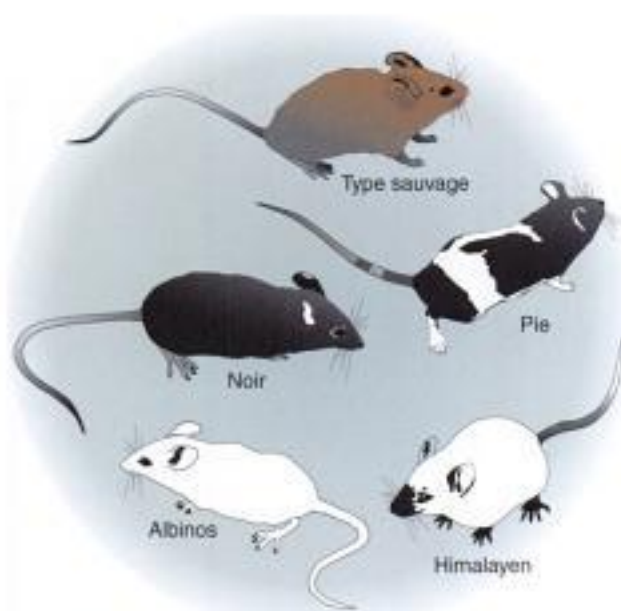


Figure 6-17 Quelques phénotypes de pelage chez les souris.

De telles interactions se déduisent de proportions modifiées dans des croisements de dihybrides. La figure 6-17 illustre quelques types de pigmentation chez les souris. Les interactions de ces types de gènes déterminent la plupart des caractères chez n'importe quel organisme.

MESSAGE

Différentes sortes de proportions modifiées de dihybrides indiquent que les gènes peuvent interagir de plusieurs façons les uns avec les autres pour déterminer un phénotype.

LA PÉNÉTRANCE ET L'EXPRESSIVITÉ

Dans les exemples précédents, l'origine génétique de la dépendance d'un gène sur un autre a été déduite de rapports génétiques précis. Cependant, seule une proportion relativement faible des gènes d'un génome présentent des propriétés qui les rendent favorables à une telle analyse. L'une des propriétés importantes est que la mutation ait une viabilité ou une fécondité égale à celle du type sauvage, de sorte que la fréquence d'obtention de classes de mutants ou de types sauvages ne soit pas modifiée.

Une autre propriété est que la différence dans la norme de réaction (voir chapitre 1) entre mutant et type sauvage soit si importante qu'il n'y ait pas de chevauchement de leurs courbes de réaction, ce qui permet de distinguer les génotypes mutants des génotypes sauvages avec une certitude de 100 pour cent. Dans ce cas, nous disons que cette mutation est pénétrante à 100 pour cent. Cependant, de nombreuses mutations présentent une pénétrance incomplète. La **pénétrance** est donc définie comme le pourcentage d'individus avec un génotype donné, qui présentent le phénotype associé à ce génotype. Par exemple, un organisme peut avoir un génotype particulier mais ne pas exprimer le phénotype correspondant en raison de gènes modificateurs épistatiques ou suppresseurs dans le reste du génome ou en raison d'un effet modificateur de l'environnement. Ou bien, l'absence de fonction d'un gène peut avoir intrinsèquement des effets très subtils, difficiles à mesurer lors d'une étude en laboratoire.

L'**expressivité** est un autre terme qui sert à décrire la gamme d'expression phénotypique. L'expressivité mesure le taux auquel un génotype donné est exprimé au niveau phénotypique. Des degrés variables d'expression chez des individus distincts peuvent être dus à une variation des allèles dans le reste du génome ou à des facteurs environnementaux. La figure 6-18 illustre la différence entre pénétrance et expressivité. Comme la pénétrance, l'expressivité fait partie intégrante du concept de norme de réaction.

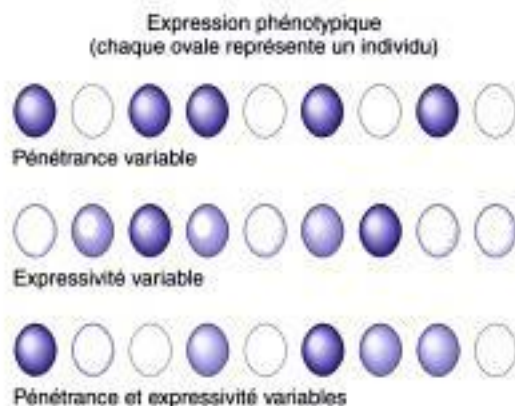


Figure 6-18 Schéma représentant les effets de la pénétrance et de l'expressivité en utilisant un caractère hypothétique « intensité du pigment ». Dans chaque ligne, tous les individus ont le même allèle, appelons-le *P*, leur conférant le même « potentiel pour produire des pigments ». Cependant, des effets dus au reste du génome et à l'environnement peuvent supprimer ou modifier la production de pigment chez un individu.

Toutes sortes d'analyses génétiques, telles que l'analyse d'un arbre généalogique humain ou les prédictions d'un conseiller génétique peuvent être fortement compliquées par les phénomènes de pénétrance incomplète et d'expressivité variable. Par exemple, si un allèle responsable d'une maladie ne présente pas une pénétrance complète (comme c'est souvent le cas), il est difficile de donner un bulletin de santé génétique exact à tous les individus d'un arbre généalogique (par exemple l'individu R dans la figure 6-19). D'autre part, l'analyse d'un arbre généalogique peut parfois permettre d'identifier des individus qui n'expriment pas, mais qui ont certainement le génotype de la maladie (par exemple, l'individu Q dans la figure 6-19). De la même façon, l'expressivité variable peut induire des erreurs dans le diagnostic. Un exemple spécifique d'expressivité variable est présenté dans la figure 6-20.

MESSAGE

Seul un petit groupe de mutations présente une pénétrance complète et une expressivité qui ne varie pas. Ces mutations, expérimentalement reproductibles, sont les plus adéquates pour tester les interactions entre les gènes.

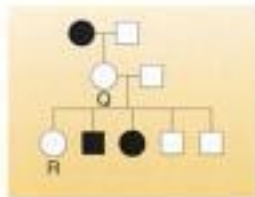


Figure 6-19 Le phénomène de pénétrance incomplète illustré par un arbre généalogique concernant un allèle dominant. L'individu Q doit avoir l'allèle (car il l'a transmis à sa descendance) mais il n'a pas été exprimé dans son phénotype. Un individu tel que R ne peut être sûr de ne pas avoir cet allèle dans son génotype.

RÉSUMÉ

Bien que par analyse génétique, il soit possible d'identifier un gène unique dont les allèles dictent deux phénotypes alternatifs pour un caractère, ce gène ne contrôle pas le caractère par lui-même : le gène doit interagir avec de nombreux autres gènes du génome. Il existe deux

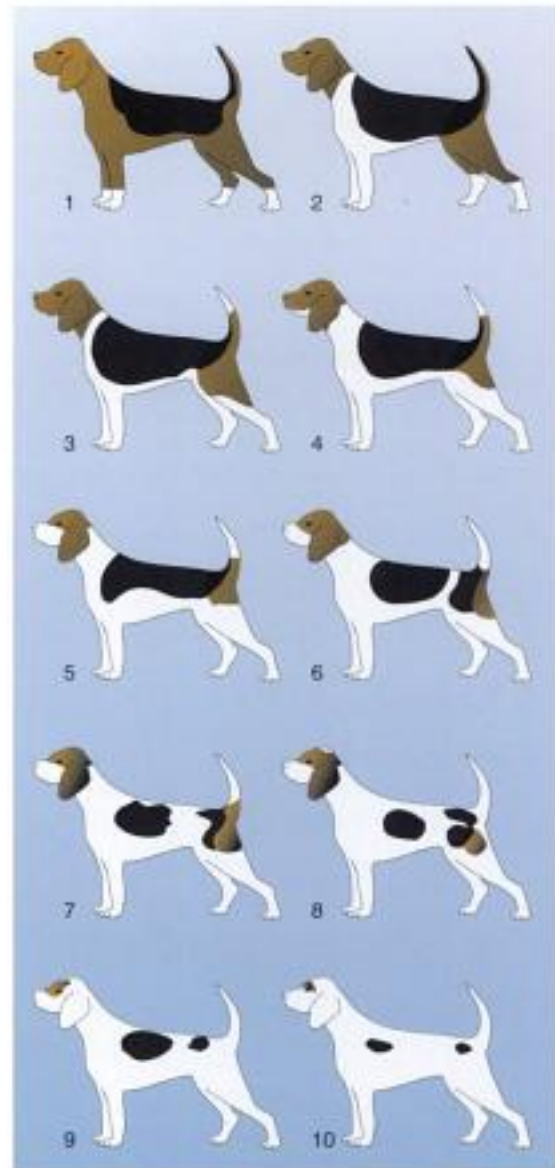


Figure 6-20 L'expressivité variable illustrée par 10 degrés de taches pie chez les beagles. Chacun des chiens possède S^p , l'allèle responsable des taches pie chez les chiens. (Adapté de Clarence C. Little. *The inheritance of coat color in dogs*, Cornell University Press, 1957 et de Giorgio Schreiber, *Journal of heredity* 9, 1930, 403).

types d'interactions : entre les allèles d'un même gène et entre des allèles de gènes différents. Le test de complémentation permet de décider si deux phénotypes mutants apparentés déterminés par des allèles récessifs sont dus à des mutations dans un même gène ou dans deux gènes

distincts. Les génotypes mutants sont réunis chez un individu de la F_1 , et si le phénotype est mutant, alors il n'y a pas eu de complémentation et les deux allèles doivent correspondre au même gène. Si on observe une complémentation, les allèles doivent correspondre à des gènes différents.

Les allèles d'un même gène interagissent de plusieurs façons. Premièrement, leurs actions cellulaires spécifiques entraînent des variations de la dominance : une dominance incomplète résulte de quantités intermédiaires de la protéine chez l'hétérozygote et la codominance est due à la production de deux produits protéiques distincts. Les allèles interagissent également par une sorte d'effet de dosage : un allèle mutant dans l'état hétérozygote peut produire un effet légèrement nuisible, alors que deux doses chez un mutant homozygote peuvent avoir des conséquences graves, voire létales.

L'interaction des allèles de différents gènes peut être détectée grâce aux rapports modifiés de dihybrides 9 : 3 : 3 : 1.

Des mutations dans des gènes appartenant à la même voie biochimique aboutissent à un rapport 9 : 7 si les phénotypes mutants sont les mêmes, ou à un rapport 9 : 3 : 4 si les phénotypes mutants sont différents. On dit que le rapport 9 : 3 : 4 résulte d'une épistasie récessive, une situation dans laquelle la mutation dans la protéine qui agit en premier dans la séquence d'événements, empêche l'expression d'un gène dont la protéine agit plus tard dans la même séquence.

La suppression est une autre forme d'interaction de deux gènes : les mutations suppressives convertissent un phénotype mutant dû à un allèle mutant d'un autre locus, en un phénotype sauvage.

La pénétrance est la proportion d'individus chez lesquels une mutation est exprimée et l'expressivité est l'importance du phénotype attribué à un allèle. La pénétrance et l'expressivité variables sont causées par des interactions avec d'autres gènes et avec l'environnement.

DIAGRAMME CONCEPTUEL

Dessinez un diagramme conceptuel mettant en relation le maximum de termes parmi ceux qui sont proposés. Remarquez que ces termes ne sont pas écrits dans un ordre particulier.

dihybride / rapport 9 : 3 : 4 / assortiment indépendant / voie biochimique / épistasie récessive / interaction de gènes / méiose / type sauvage

PROBLÈME RÉSOLU

1. Les scarabées d'une espèce particulière peuvent avoir des élytres de couleur verte (*green*, en anglais), bleue ou turquoise. Des scarabées vierges ont été sélectionnés dans une population polymorphe de laboratoire et croisés pour déterminer le mode de transmission de la couleur des élytres. Les croisements et les résultats ont été les suivants :

Croisement	Parents	Descendants
1	Bleu × vert	Tous bleu
2	Bleu × bleu	$\frac{3}{4}$ bleu : $\frac{1}{4}$ turquoise
3	Vert × vert	$\frac{3}{4}$ vert : $\frac{1}{4}$ turquoise
4	Bleu × turquoise	$\frac{1}{2}$ bleu : $\frac{1}{2}$ turquoise
5	Bleu × bleu	$\frac{3}{4}$ bleu : $\frac{1}{4}$ vert
6	Bleu × vert	$\frac{1}{2}$ bleu : $\frac{1}{2}$ vert
7	Bleu × vert	$\frac{3}{4}$ bleu : $\frac{1}{4}$ vert $\frac{1}{4}$ turquoise
8	Turquoise × turquoise	Tous turquoise

- a. Déduisez l'explication génétique de la couleur des élytres dans cette espèce.
- b. Écrivez de la façon la plus précise possible, les génotypes des parents et des descendants.

SOLUTION

- a. Ces données semblent complexes à première vue, mais le mode de transmission devient clair si l'on considère les croisements les uns après les autres. Pour résoudre ce type de problème, il faut commencer par regarder tous les croisements et par regrouper les données pour en déduire les modes de transmission. Un élément ressort de la considération globale des données : tous les rapports correspondent à ceux d'un gène unique. Il n'y a pas de preuve que deux gènes distincts soient impliqués. Comment une telle variation peut-elle s'expliquer avec un seul gène ? La réponse évidente est qu'il y a variation pour le gène lui-même — c'est-à-dire un allélisme multiple. Il y a sans doute trois allèles du même gène. Appelons le gène w et désignons les allèles par w^b , w^t et w^v . Nous avons maintenant un problème supplémentaire, qui est de déterminer les relations de dominance entre ces allèles.

Le croisement 1 nous apprend quelque chose sur la dominance, car les descendants d'un croisement bleu x vert sont tous bleus ; le bleu semble donc dominant sur le vert. Cette conclusion est renforcée par le croisement 5, car le déterminant vert devait être présent dans la souche parentale pour apparaître dans la descendance. Le croisement 3 nous donne des renseignements sur le déterminant turquoise, qui devait être présent, bien qu'inexprimé, dans la souche parentale car certains des descendants de ce croisement présentent des élytres turquoises. Le vert doit donc être dominant sur le turquoise. Nous

venons d'établir un modèle dans lequel la dominance est $w^b > w^t > w^v$. De plus, la position en fin de cette série allélique, déduite pour w^v , est confirmée par les résultats du croisement 7, dans lequel la couleur turquoise apparaît dans la descendance d'un croisement bleu x vert.

- b. Il s'agit maintenant seulement de déduire les génotypes spécifiques. Remarquez que l'énoncé précise que les parents ont été retirés d'une population polymorphe ; ceci signifie qu'ils pouvaient être soit homozygotes, soit hétérozygotes. Un parent avec des élytres bleues par exemple, pouvait être homozygote (w^b/w^b) ou hétérozygote (w^b/w^t ou w^b/w^v). Ici un peu d'essai-erreur et de bon sens sont nécessaires mais à ce stade, la question a presque été résolue et il reste seulement à « mettre les points sur les i ». Les génotypes suivants expliquent les résultats. Un tiret indique que le génotype peut être soit homozygote, soit hétérozygote s'il comporte un allèle situé plus « bas » dans la série allélique classée par ordre décroissant de dominance.

Croisement	Parents	Descendants
1	$w^b/w^b \times w^v/w^v$	w^b/w^t ou w^b/w^v
2	$w^b/w^t \times w^b/w^t$	$\frac{3}{4} w^b/w^b$: $\frac{1}{4} w^t/w^t$
3	$w^b/w^t \times w^t/w^t$	$\frac{3}{4} w^b/w^t$: $\frac{1}{4} w^t/w^t$
4	$w^b/w^t \times w^v/w^v$	$\frac{1}{2} w^b/w^v$: $\frac{1}{2} w^t/w^v$
5	$w^b/w^t \times w^v/w^v$	$\frac{3}{4} w^b/w^v$: $\frac{1}{4} w^t/w^v$
6	$w^b/w^t \times w^t/w^t$	$\frac{1}{2} w^b/w^t$: $\frac{1}{2} w^t/w^t$
7	$w^b/w^t \times w^t/w^v$	$\frac{1}{2} w^b/w^v$: $\frac{1}{4} w^b/w^t$: $\frac{1}{4} w^t/w^v$
8	$w^t/w^t \times w^v/w^v$	tous w^t/w^v

PROBLÈME RÉSOLU

2. Les feuilles d'ananas peuvent être classées en trois types : épineux (S), à bouts épineux (ST) et non épineux (P). Lors de croisements entre des souches pures, suivis par des croisements entre membres de la F_1 , les résultats suivants ont été obtenus :

Croisement	Parents	Phénotypes	
		F_1	F_2
1	ST \times S	ST	99 ST : 34 S
2	P \times ST	P	120 P : 39 ST
3	P \times S	P	95 P : 25 ST : 8 S

- a. Attribuez des symboles aux gènes. Expliquez ces résultats en fonction des génotypes et de leurs rapports.
- b. En utilisant le modèle de la partie a, indiquez les rapports phénotypiques que vous attendriez si vous croisez (1) la descendance F_1 issue de non épineux x épineux avec la souche parentale épineux et (2) la descendance F_1 de non épineux x épineux avec la descendance F_1 de épineux x bouts épineux.

SOLUTION

- a. Premièrement, regardons les rapports de la F_2 . Nous avons des rapports 3 : 1 évidents dans les croisements 1 et 2, ce qui indique des ségrégations d'un gène unique. Le croisement 3 au contraire, présente un rapport voisin de 12 : 3 : 1. Ceci serait une modification d'un rapport de dihybrides 9 : 3 : 3 : 1, qui nous laisse donc penser à une interaction entre deux gènes. Cela semble le meilleur point de départ. Nous reviendrons plus tard sur les croisements 1 et 2. Toute autofécondation d'un dihybride, impliquant deux gènes dont l'assortiment est indépendant (avec une dominance complète dans les deux cas) est basée sur des proportions 9 : 3 : 3 : 1. Puisque nous ne savons pas quel phénotype est sauvage, nous ne pouvons utiliser le symbolisme de type sauvage. Utilisons donc simplement le symbolisme générique A/a et B/b pour représenter les deux locus en interaction. Les génotypes doivent être regroupés de la façon suivante pour obtenir un rapport 12 : 3 : 1 :

9 $A/- ; B/-$	} 12 non épineux
3 $A/- ; b/b$	
3 $a/a ; B/-$	3 bouts épineux
1 $a/a ; b/b$	1 épineux

Sans nous inquiéter du nom du type d'interaction des gènes (on ne nous le demande pas de toute façon), nous pouvons déjà définir nos trois phénotypes de feuilles d'ananas d'après les paires alléliques proposées, A , a et B , b :

Non épineux	$A/-$ (B , b indifférent)
Bouts épineux	$a/a ; B/-$
Épineux	$a/a ; b/b$

Remarque : Nous aurions pu les désigner en inversant les lettres A et B : $B/-$ (A ou a , indifférent), $b/b ; A/-$ et $b/b ; a/a$; cela ne fait aucune différence).

Qu'en est-il des parents du croisement 3 ? Le parent épineux doit être $a/a ; b/b$, et comme le gène B est nécessaire pour produire les individus à bouts épineux de la F_2 , le parent non épineux doit être $A/A ; B/B$. (Remarquez que l'on nous dit que tous les parents sont de lignée pure, ou homozygotes). La F_1 doit donc être $A/a ; B/b$. Sans plus attendre, nous pouvons écrire le croisement 1 de la façon suivante :

$$a/a ; B/B \times a/a ; b/b \longrightarrow a/a ; B/b \begin{cases} \frac{3}{4} a/a ; B/b \\ \frac{1}{4} a/a ; b/b \end{cases}$$

Le croisement 2 peut également être écrit presque avec certitude, en utilisant nos symboles arbitraires de gènes :

$$A/A ; -/- \times a/a ; B/B \longrightarrow A/a ; B/- \begin{cases} \frac{1}{4} A/- ; -/- \\ \frac{3}{4} a/a ; B/- \end{cases}$$

Nous savons que la F_2 du croisement 2 présente une ségrégation correspondant à un gène unique, et il semble certain à présent que la paire allélique A/a soit impliquée. Mais l'allèle B est nécessaire pour produire le phénotype à bouts épineux. Tous les individus doivent donc être homozygotes B/B :

$$A/A ; B/B \times a/a ; B/B \longrightarrow A/a ; B/B \begin{cases} \frac{1}{4} A/- ; B/B \\ \frac{3}{4} a/a ; B/B \end{cases}$$

Remarquez que les deux ségrégations de gène unique dans les croisements 1 et 2 ne signifient pas que les gènes n'interagissent pas. On voit que l'interaction des deux gènes n'est pas révélée par ces croisements — seulement par le 3, dans lequel la F_1 est hétérozygote pour les deux gènes.

- b. Il suffit maintenant d'utiliser les lois de la ségrégation et de l'assortiment indépendant pour prédire les résultats des croisements :

$$1) \quad A/a ; B/b \times a/a ; b/b \longrightarrow \begin{cases} \frac{1}{2} A/a ; B/b \\ \frac{1}{2} A/a ; b/b \end{cases} \left. \begin{array}{l} \text{(assortiment} \\ \text{indépendant} \\ \text{dans un croisement-} \\ \text{test standard)} \end{array} \right\} \begin{array}{l} \text{non épineux} \\ \frac{1}{2} a/a ; B/b \text{ bouts épineux} \\ \frac{1}{2} a/a ; b/b \text{ épineux} \end{array}$$

$$2) \quad A/a ; B/b \times a/a ; B/b \longrightarrow \begin{array}{l} \frac{1}{2} A/a \begin{cases} \frac{3}{4} B/- \longrightarrow \frac{3}{8} \\ \frac{1}{4} b/b \longrightarrow \frac{1}{8} \end{cases} \left. \begin{array}{l} \\ \end{array} \right\} \frac{1}{2} \text{ non épineux} \\ \frac{1}{2} a/a \begin{cases} \frac{3}{4} B/- \longrightarrow \frac{3}{8} \text{ bouts épineux} \\ \frac{1}{4} b/b \longrightarrow \frac{1}{8} \text{ épineux} \end{cases} \end{array}$$

PROBLÈMES

1. La recombinaison et la complémentation peuvent toutes deux partir de deux mutants récessifs différents et aboutir à un type sauvage. Comment expliqueriez-vous la différence entre ces deux processus ?
2. Les poules herminettes ont des plumes claires pour la plupart, avec occasionnellement une plume noire, ce qui leur donne une apparence tachetée. Un croisement entre deux herminettes a produit un total de 48 descendants, comprenant 22 herminettes, 14 noires et 12 blanches. Quelle explication génétique du motif du plumage des herminettes est suggérée ici ? Comment pourriez-vous tester ces hypothèses ?
3. Dans la série d'allèles multiples qui détermine la couleur du pelage des lapins, c^+ code l'aspect agouti, c^b , le chinchilla et c^h , l'himalayen. La dominance existe dans l'ordre suivant $c^+ > c^b > c^h$. Dans un croisement $c^+ / c^b \times c^b / c^h$, quelle proportion de descendants sera chinchilla ?
4. Noir (*black* en anglais), sépia, crème et albinos sont des couleurs de pelage de cobayes. Des animaux (qui n'étaient pas nécessairement de lignée pure) de ces couleurs ont été croisés les uns avec les autres. Les résultats sont regroupés dans le tableau ci-dessous, où l'abréviation A correspond à albinos, B à noir, C à crème et S à sépia, pour les phénotypes :

Croisement	Phéno- types parentaux	Phénotypes des descendants			
		B	S	C	A
1	B × B	22	0	0	7
2	B × A	10	9	0	0
3	C × C	0	0	34	11
4	S × C	0	24	11	12
5	B × A	13	0	12	0
6	B × C	19	20	0	0
7	B × S	18	20	0	0
8	B × S	14	8	6	0
9	S × S	0	26	9	0
10	C × A	0	0	15	17

- a. Déduisez le mode de transmission de ces couleurs de pelage, en utilisant les symboles de gènes de votre choix. Indiquez les génotypes de tous les parents et de tous les descendants.
- b. Si les animaux noirs des croisements 7 et 8 sont croisés, quelles proportions pouvez-vous prédire parmi les descendants, en utilisant votre modèle ?

5. Dans une maternité, quatre bébés ont été accidentellement mélangés. On sait que les types sanguins ABO des quatre bébés sont O, A, B et AB. Les quatre types sanguins des quatre couples de parents ont été déterminés. Indiquez le bébé qui correspond à chaque couple de parents : (a) AB × O, (b) A × O, (c) A × AB, (d) O × O.
6. Deux mouches du vinaigre de phénotype sauvage ont été croisées et la descendance comprenait 202 femelles et 98 mâles.
 - a. Qu'y a-t-il d'inhabituel dans ce résultat ?
 - b. Donnez une explication génétique à cela.
 - c. Indiquez un test de votre hypothèse.
7. Une lignée pure de gourdes (courges) qui produisent des fruits en forme de disque (voir la figure ci-dessous) a été croisée avec une lignée pure à longs fruits. La F_1 présente des fruits en forme de disque, mais la F_2 présente un phénotype nouveau, sphérique, composé des proportions suivantes :

Disque	270
Sphère	178
Long	32



Proposez une explication pour ces résultats et donnez les génotypes de la génération parentale P et des générations F_1 et F_2 .

(Illustration d'après P. J. Russell, *Genetics*, 3^e édition. Harper-Collins, 1992).

8. Les gueules-de-loup de type sauvage (*Antirrhinum*) synthétisent un pigment anthocyanine rouge et ont des pétales rouges. Deux lignées mutantes pures d'*Antirrhinum* dépourvues d'anthocyanine

ont été développées, l'une en Californie et l'autre en Hollande. Elles semblaient identiques en raison de leur absence totale de pigment rouge, qui se manifestait par des fleurs blanches (albinos). Cependant, lorsque les pétales des deux lignées ont été broyés ensemble dans un tampon, à l'intérieur d'un tube à essai, la solution, qui semblait incolore au début, devint progressivement rouge.

- Quelles expériences de contrôle un chercheur devrait-il réaliser avant de continuer son analyse ?
- Qu'est-ce qui pourrait expliquer la formation de la couleur rouge dans le tube à essai ?
- D'après votre explication pour la partie b, quels seraient les génotypes des deux lignées ?
- Si les deux lignées blanches étaient croisées, quels phénotypes prédiriez-vous pour la F_1 et la F_2 ? Quel principe génétique est illustré par les résultats de la F_1 ? Par les résultats de la F_2 ?

9. Quatre lignées mutantes homozygotes récessives de *Drosophila melanogaster* (marquées de 1 à 4) présentent une coordination anormale des pattes, ce qui les fait marcher de manière très irrégulière. Ces lignées sont croisées les unes avec les autres. Les phénotypes des mouches de la F_1 sont indiqués dans le tableau suivant, où « + » signifie que la F_1 présente une marche de type sauvage et où « - » signifie une marche anormale.

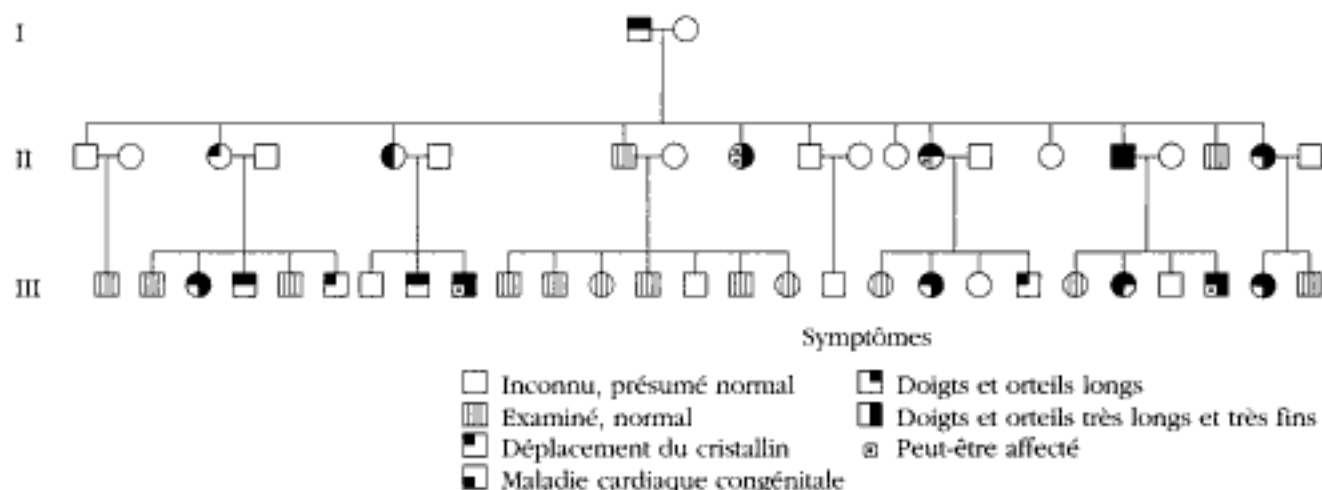
	1	2	3	4
1	-	+	+	+
2	+	-	-	+
3	+	-	+	+
4	+	+	+	+

- Quel type de test cette analyse représente-t-elle ?
- Combien de gènes différents ont été mutés en créant ces quatre lignées ?
- Inventez des symboles pour les allèles mutants et sauvage et écrivez les génotypes complets des quatre lignées et de la F_1 .
- Ces données nous disent-elles quels gènes sont liés ? Si ce n'est pas le cas, comment pourrait-on tester la liaison génétique ?
- Ces données nous indiquent-elles le nombre total de gènes impliqués dans la coordination des pattes chez cet animal ?

10. Trois mutants de levure haploïde, auxotrophes pour le tryptophane (qui ont besoin de tryptophane pour croître) et isolés indépendamment, sont appelés *trpB*, *trpD* et *trpE*. Des suspensions cellulaires de chacun ont été étalées sur une boîte de culture contenant un milieu nutritif enrichi avec juste assez de tryptophane pour permettre une faible croissance d'une souche *trp*. Les trainées sont disposées en un motif triangulaire, de façon à ce qu'elles ne se touchent pas. On note une croissance très importante aux deux extrémités de la bande *trpE* et à une extrémité de la bande *trpD* (voir la figure ci-dessous).



- Pensez-vous que la complémentation intervienne ici ?
 - Expliquez brièvement la disposition des zones de croissance abondante.
 - Quel est l'ordre des étapes enzymatiques déficientes chez *trpB*, *trpD* et *trpE* dans la voie de synthèse du tryptophane ?
 - Pourquoi était-il nécessaire d'ajouter une petite quantité de tryptophane au milieu, afin de démontrer un tel mode de croissance ?
11. La poule frisée est très admirée des amateurs de volailles. Elle tire son nom de la façon inhabituelle dont ses plumes se recourbent, donnant l'impression (selon les mots mémorables du généticien animalier F. B. Hutt) « d'avoir été caressée à rebrousse-plumes ». Malheureusement il n'existe pas de lignée pure de poules frisées. Lorsque deux poules frisées sont croisées, elles produisent toujours 50 pour cent de frisées, 25 pour cent de normales et 25 pour cent avec des plumes laineuses particulières qui tombent rapidement, laissant les volailles nues.
- Donnez une explication génétique détaillée de ces résultats, en indiquant les génotypes de tous les phénotypes.
 - Si vous vouliez réaliser une production de masse des poules frisées pour les vendre, quels types utiliseriez-vous de préférence pour votre paire de reproducteurs ?



12. Le syndrome de Marfan est une maladie du tissu conjonctif fibreux qui se caractérise par de nombreux symptômes, dont des doigts et des orteils longs et fins, des défauts de l'œil, une maladie cardiaque et de longs membres. (Flo Hyman, la star américaine de volley-ball souffrait du syndrome de Marfan. Elle mourut peu après un match d'une rupture de l'aorte).

a. Utilisez l'arbre généalogique ci-dessus pour proposer un mode de transmission du syndrome de Marfan.

b. Quel phénomène génétique apparaît dans cet arbre ?

c. Imaginez une raison à un tel phénomène.

(Illustration d'après J. V. Neel et W. J. Schull, *Human heredity*. University of Chicago Press, 1954).

13. Une femme possédait un caniche femelle albinos de lignée pure (un phénotype récessif autosomique) et voulait l'accoupler pour produire une portée de chiots blancs. Elle emmena la chienne chez un éleveur qui affirma qu'il allait la croiser avec un mâle reproducteur albinos, également de lignée pure. Six chiots naquirent, tous noirs et la femme poursuivit l'éleveur en justice, affirmant qu'il avait remplacé le mâle reproducteur par un chien noir, donnant ces six chiots non désirés. Vous êtes appelé comme témoin expert et la défense vous demande s'il est possible de produire des descendants noirs à partir de deux parents albinos récessifs de lignées pures. Quel témoignage donnez-vous ?

14. Les pétales de la plante *Collinsia parviflora* sont normalement bleus. Deux lignées pures ont été obtenues à partir de variants de couleur découverts dans la nature ; la première lignée avait des pétales roses et la deuxième, des pétales blancs. Les croisements suivants ont été réalisés entre les lignées pures, avec les résultats indiqués ci-dessous :

Parents	F ₁	F ₂
Bleu × blanc	Bleu	101 bleu, 33 blanc
Bleu × rose	Bleu	192 bleu, 63 rose
Rose × blanc	Bleu	272 bleu, 121 blanc, 89 rose

a. Expliquez ces résultats génétiquement. Définissez les symboles d'allèles que vous utilisez et indiquez la constitution génétique des parents, de la F₁ et de la F₂.

b. Un croisement entre une certaine plante bleue de la F₂ et une certaine plante blanche de la F₂ a donné des descendants dont $\frac{3}{8}$ étaient bleus, $\frac{1}{8}$ étaient roses et $\frac{1}{2}$ étaient blancs. Quels devaient être les génotypes de ces deux plantes de la F₂ ?

Décomposons le problème

a. Quel est le caractère étudié ?

b. Quel est le phénotype sauvage ?

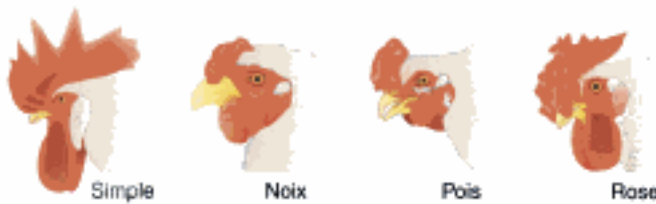
c. Qu'est-ce qu'un variant ?

- d. Quels sont les variants dans ce problème ?
- e. Que signifie « dans la nature » ?
- f. De quelle façon les variants ont-ils pu être trouvés dans la nature ? (Décrivez la scène).
- g. À quelles étapes des expériences les graines ont-elles été utilisées ?
- h. Écrire un croisement « bleu x blanc » (par exemple) signifie-t-il la même chose qu'écrire « blanc x bleu » ? Vous attendez-vous à des résultats identiques ? Pourquoi ou pourquoi non ?
- i. En quoi les deux premières lignes du tableau sont-elles différentes de la troisième ?
- j. Quels sont les phénotypes dominants ?
- k. Qu'est-ce que la complémentation ?
- l. D'où provient la couleur bleue chez les descendants du croisement rose x blanc ?
- m. Quel phénomène génétique représente la production d'une F_1 bleue à partir de parents roses et blancs ?
- n. Indiquez le plus de rapports possibles.
- o. Y a-t-il des rapports correspondant à des monohybrides ?
- p. Y a-t-il des rapports correspondant à des dihybrides ?
- q. Que nous apprend l'observation des rapports des mono- et dihybrides ?
- r. Indiquez quatre rapports mendéliens modifiés qui vous viennent à l'esprit.
- s. Y a-t-il des rapports mendéliens modifiés dans ce problème ?
- t. Qu'indiquent généralement des rapports mendéliens modifiés ?
- u. Qu'indiquent le ou les rapports spécifiques modifiés dans ce problème ?
- v. Dessinez les chromosomes lors de la méiose chez les parents du croisement bleu x blanc et lors de la méiose de la F_1 .
- w. Faites la même chose pour le croisement bleu x rose.

15. La plupart des coléoptères de la farine sont noirs, mais on connaît plusieurs variants de la couleur. Des croisements entre parents de lignée pure ont donné les résultats suivants dans la génération F_1 et le croisement de membres de la F_1 issus de chaque croisement, a donné les rapports indiqués pour la génération F_2 . Les phénotypes sont abrégés en Bl, noir (*black* en anglais) ; Br, marron (*brown* en anglais) ; Y, jaune (*yellow* en anglais) et W, blanc (*white* en anglais).

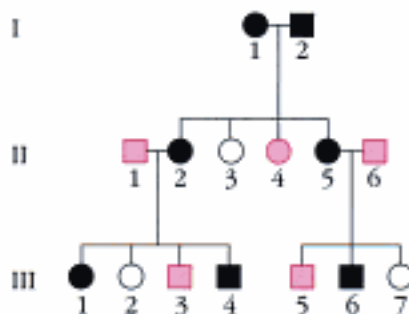
Croisement	Parents	F_1	F_2
1	Br × Y	Br	3 Br:1 Y
2	Bl × Br	Bl	3 Bl:1 Br
3	Bl × Y	Bl	3 Bl:1 Y
4	W × Y	Bl	9 Bl:3 Y:4 W
5	W × Br	Bl	9 Bl:3 Br:4 W
6	Bl × W	Bl	9 Bl:3 Y:4 W

- a. D'après ces résultats, déduisez et expliquez la transmission de ces couleurs.
 - b. Écrivez les génotypes complets de chaque parent, de la F_1 et de la F_2 dans tous les croisements.
 - c. Ordonnez les gènes et les pigments en une ou plusieurs voies biochimiques.
16. Un allèle A , qui n'est pas létal à l'état homozygote, donne aux rats qui le possèdent un pelage jaune. L'allèle R d'un gène séparé, dont l'assortiment est indépendant du premier, produit un pelage noir. A et R ensemble produisent un pelage gris, alors que a et r produisent un pelage blanc. Un mâle gris est croisé avec une femelle jaune et la F_1 comporte 3/8 de jaunes, 3/8 de gris, 1/8 de noirs et 1/8 de blancs. Déterminez les génotypes des parents.
 17. Chez la poule, le génotype r/r ; p/p donne une crête de type simple, $R/-$; $P/-$ une crête de type noix (découpée comme une noix), r/r ; $P/-$ une crête de type pois et $R/-$; p/p une crête de type rose (voir les illustrations). Les gènes ne sont pas liés.



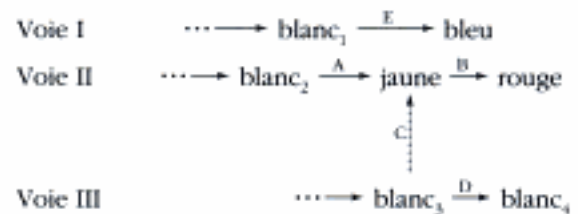
- Quels types de crête apparaîtront dans la F_1 et dans la F_2 et dans quelles proportions si l'on croise des volailles à crête simple avec des volailles de lignée pure à crête de type noix ?
- Quels sont les génotypes des parents dans un croisement noix x rose dont les descendants sont $3/8$ rose, $3/8$ noix, $1/8$ pois et $1/8$ simple ?
- Quels sont les génotypes des parents d'un croisement noix x rose dont toute la descendance est de type noix ?
- Combien de génotypes produisent un phénotype noix ? Écrivez-les tous.

18. L'allèle B donne des souris à pelage noir (*black* en anglais) et b des souris à pelage marron (*brown* en anglais). Le génotype e/e correspondant à un autre gène, dont l'assortiment est indépendant du précédent, empêche l'expression de B et b , donnant un pelage beige, tandis que $E/-$ permet l'expression de B et b . Les deux gènes sont autosomiques. Dans l'arbre généalogique suivant, les symboles noirs indiquent un pelage noir, les symboles roses un pelage marron et les symboles blancs un pelage beige :



- Quel est le nom donné au type d'interaction des gènes dans cet exemple ?
- Quels sont les génotypes des individus de cet arbre ? (S'il y a plusieurs possibilités, citez-les).

19. Supposez que deux pigments, rouge et bleu, s'additionnent pour donner la couleur violette normale de type sauvage aux pétales de pétunia. Des voies biochimiques séparées synthétisent les deux pigments, comme le montrent les deux lignes du haut dans la figure ci-dessous. • Blanc • fait référence à des composés qui ne sont pas des pigments. (L'absence totale de pigment conduit à des pétales blancs). Le pigment rouge se forme à partir d'un intermédiaire jaune dont la concentration est trop faible en temps normal pour colorer les pétales. Une troisième voie dont les composants ne contribuent pas à la pigmentation des pétales n'affecte normalement pas les voies du bleu et du rouge, mais si l'on pouvait augmenter la concentration de l'un de ses intermédiaires (blanc_3), il pourrait être converti en l'intermédiaire jaune de la voie rouge. Dans le schéma, les lettres A à E représentent des enzymes ; les gènes qui leur correspondent, qui sont tous non liés génétiquement, peuvent être symbolisés par les mêmes lettres.



Supposez que les allèles de type sauvage soient dominants et codent la fonction enzymatique, et que les allèles récessifs représentent l'absence de fonction enzymatique. Déduisez les combinaisons des génotypes parentaux de lignée pure dont le croisement aurait pu produire les descendants de la F_2 dans les rapports suivants :

- 9 violet : 3 vert : 4 bleu
- 9 violet : 3 rouge : 3 bleu : 1 blanc
- 13 violet : 3 bleu
- 9 violet : 3 rouge : 3 vert : 1 jaune
- Quelle est la mutation qui agit comme un supprimeur ?
- Quelle est la mutation qui présente une épistasie récessive ?

(Remarque : Le bleu mélangé à du jaune donne du vert ; supposez qu'aucune mutation n'est létale).

Hidden page

nombre représentent l'effectif des descendants sur les 500, qui ont pu synthétiser eux-mêmes de la leucine (et ont poussé sur le milieu dépourvu de leucine) :

	a	b	c	d	e
a	0	125	128	126	0
b		0	124	2	125
c			0	124	127
d				0	123
e					0

- Dans les expériences I et II, des souches leucine-indépendantes ont été créées. Expliquez les différentes origines de ces souches dans les deux expériences.
- Expliquez la distribution des + et des - chez les hétérocaryons, pour les différents appariements de l'expérience I.
- Expliquez les différentes fréquences de souches leucine-indépendantes dans l'expérience II. (Remarque : on sait que les deux descendants leucine-indépendants issus du croisement $b \times d$ ne sont pas dus à une réversion vers le type sauvage.)
- Faites un schéma qui résume les résultats des deux expériences.

24. Le champignon *Neurospora* (un champignon haploïde) de type sauvage peut synthétiser sa propre adénine à partir de composants inorganiques contenus dans le milieu de croissance. Lors d'une analyse mutationnelle de la voie de synthèse de l'adénine, 10 mutations entraînant un besoin en arginine chez *Neurospora* ont été obtenues et des hétérocaryons ont été fabriqués en combinant les souches deux à deux. Les résultats sont présentés dans le tableau suivant, dans lequel un signe ++ signifie que l'hétérocaryon s'est développé et un signe - qu'il ne s'est pas développé :

	1	2	3	4	5	6	7	8	9	10
1	-	-	+	+	+	+	-	+	+	+
2		-	+	+	+	+	-	+	+	+
3			-	+	-	+	+	-	+	+
4				-	+	-	+	+	-	-
5					-	+	+	-	+	+
6						-	+	+	-	-
7							-	+	+	+
8								-	+	+
9									-	-
10										-

- Combien de gènes sont impliqués dans ces mutations ?

On teste la croissance des mutants 1, 3 et 4 sur les composés CAIR, AIR et SAICAR, tous chimiquement apparentés à l'adénine. Les résultats sont les suivants, où + signifie qu'il y a croissance, et - qu'il n'y a pas de croissance :

	Adénine	CAIR	AIR	SAICAR
1	+	-	-	+
3	+	-	-	-
4	+	+	-	+

- Expliquez ces résultats.

Les mutants 1, 3 et 4 sont croisés les uns avec les autres et 1000 ascospores sont étalées à partir de chaque croisement, sur un milieu dépourvu d'adénine, pour sélectionner les types sauvages. Les résultats sont les suivants :

	1	3	4
1	0	4	245
3		7	260
4			255

- Expliquez ces résultats et intégrez-les aux résultats précédents pour donner une brève explication de ce système génétique.

25. Chez un végétal, le gène e^+ code une certaine enzyme. Le gène r^+ code une protéine régulatrice, qui doit se fixer à la région régulatrice du gène e^+ pour déclencher la transcription. Une mutation complète récessive est obtenue dans chacun des gènes, r et e , respectivement. On fait subir une autofécondation à un dihybride $e^+/e^-; r^+/r^-$. Dans la descendance, quelle sera la proportion d'individus avec cette enzyme ? (On suppose que l'absence de cette enzyme n'est pas létale et que les gènes sont situés sur des chromosomes différents.)

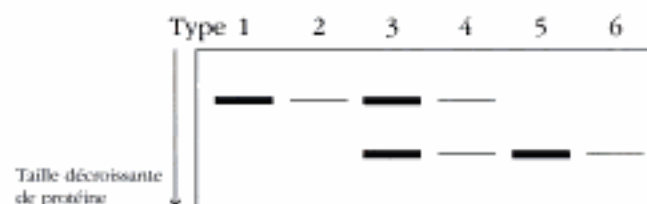
26. On sait qu'une certaine protéine de drosophile (appelons-la P) est nécessaire pour un vol normal. On sait que cette protéine est composée d'une seule chaîne polypeptidique. Deux lignées pures de mutants (lignées 1 et 2), incapables de voler, ont été obtenues. Lorsqu'on les croise avec le type sauvage, elles donnent toutes deux le même résultat :

Non volante \times type sauvage $\longrightarrow F_1$ de type sauvage $\longrightarrow F_2 \frac{3}{4}$ type sauvage : $\frac{1}{4}$ non volante

Lorsque les deux lignées pures sont croisées entre elles, les résultats sont :

Non volante 1 \times non volante 2 \longrightarrow F_1 toutes de type sauvage \longrightarrow F_2 $\frac{9}{16}$ type sauvage; $\frac{7}{16}$ non volante

On fait migrer la protéine P produite par *chacun* des membres de la F_2 dans un gel d'électrophorèse et on la marque à l'aide d'un anticorps spécifique de la protéine P. On découvre que les individus de la F_2 peuvent être regroupés en six types différents comme suit :



(Les bandes minces représentent un très faible marquage par l'anticorps.)

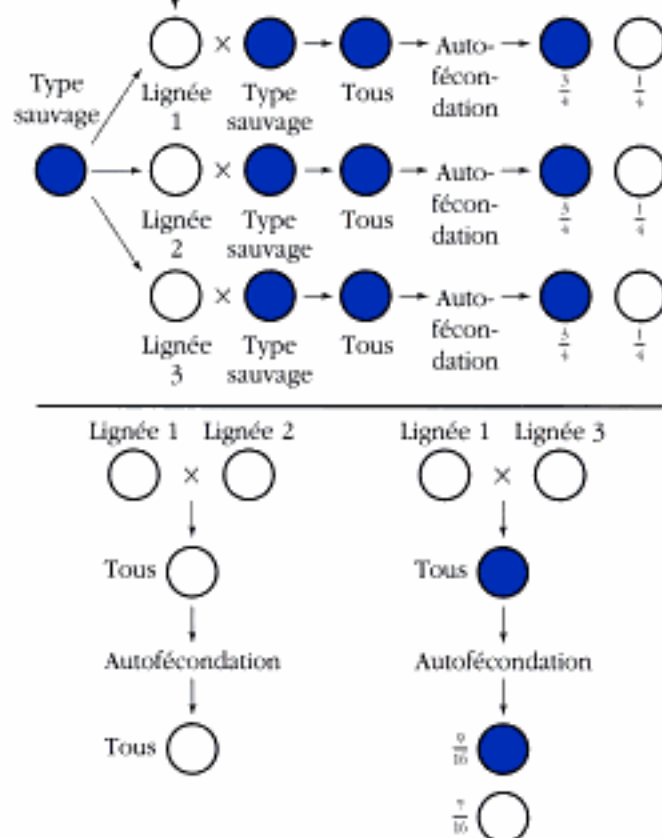
- Donnez une explication générale des résultats moléculaires révélés par l'analyse sur gel. En particulier, exposez la différence dans la nature moléculaires des deux mutations d'origine.
- D'après votre explication, quels sont les génotypes des six types de F_2 trouvés ?
- Quelle fréquence attend-on pour le type 4 ?
- Lequel ou lesquels des six types de gels de la F_2 représente les $9/16$ de phénotype sauvage ?
- Dessinez le gel auquel on s'attend pour les deux mutants parentaux et pour la F_1 .

Problèmes de reconnaissance de caractéristiques observables

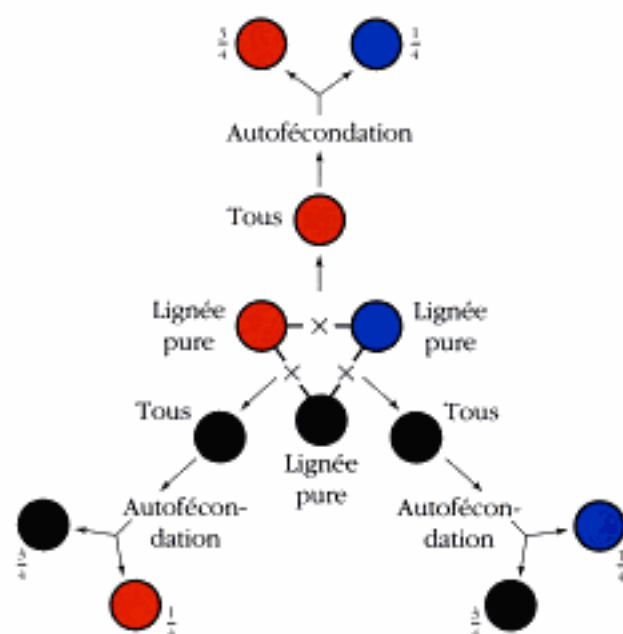
Dans les problèmes 27 à 32, les schémas montrent les phénotypes et les résultats d'analyses de croisements. Déduisez dans la mesure du possible les génotypes des individus présentés dans chaque schéma. Tous les organismes sont diploïdes.

27.

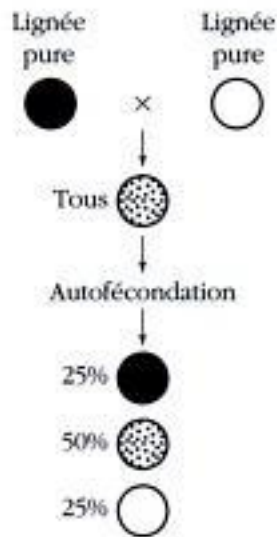
Trois lignées pures mutantes obtenues indépendamment



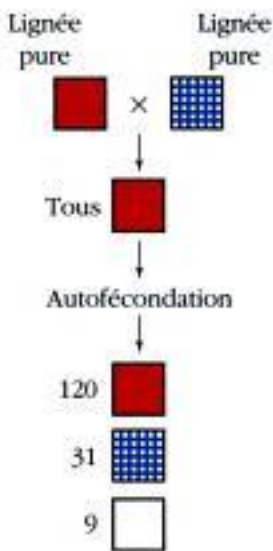
28.



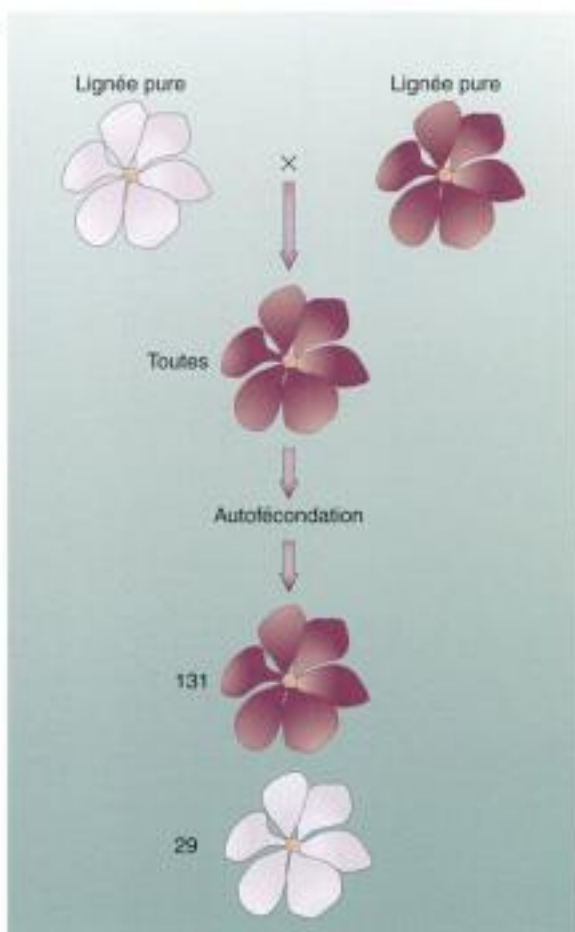
29.



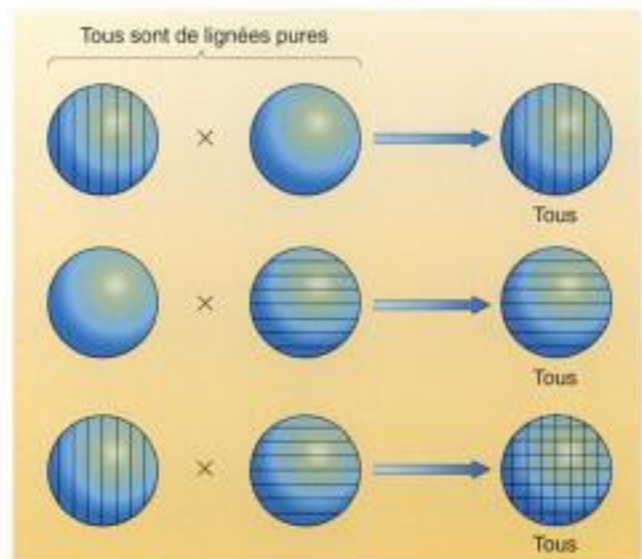
31.



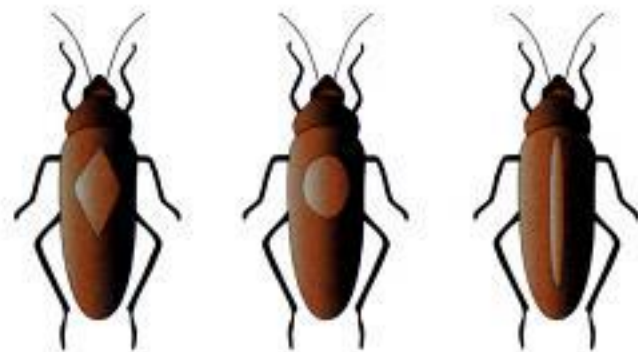
30.



32. En utilisant le schéma ci-dessous, déduisez le génotype et la dominance et prédisez les génotypes des F_2 .



33. Les élytres d'une variété de coléoptères présentent l'un des trois phénotypes suivants, un losange, un rond ou une bande. Les croisements suivants ont été réalisés entre les différentes souches qui n'étaient pas nécessairement de lignée pure :



1. Losange × losange	$\frac{3}{4}$ losange $\frac{1}{4}$ bande
2. Rond × rond	$\frac{6}{8}$ rond $\frac{1}{8}$ losange $\frac{1}{8}$ bande
3. Rond × bande	$\frac{1}{2}$ rond $\frac{1}{4}$ losange $\frac{1}{4}$ bande

- a. Inventez des symboles d'allèles, expliquez clairement l'action de tous les allèles que vous avez inventés, puis écrivez les génotypes complets des parents et des descendants des trois croisements. (Si votre explication implique plusieurs gènes, assurez-vous que *tous* les gènes sont représentés dans les génotypes de chaque croisement).
- b. Quel rapport phénotypique serait obtenu à partir du croisement entre les descendants avec un rond, issus du croisement 3 ?

7

LES MUTATIONS GÉNIQUES

POUR L'ESSENTIEL

La mutation est le processus par lequel des gènes passent d'une forme allélique à une autre.

Les mutations peuvent apparaître spontanément, à cause de différents mécanismes, dont des erreurs dans la réplication de l'ADN et des lésions spontanées de l'ADN.

Les mutagènes sont des agents qui augmentent la fréquence de mutagenèse, généralement en modifiant l'ADN.

Les systèmes de réparation biologique éliminent de nombreux changements potentiellement mutagènes dans l'ADN.

Les cellules dépourvues de certains systèmes de réparation ont des taux de mutation plus élevés que la normale.

Les mutations directes sont des changements par rapport à l'allèle de type sauvage, et les réversions sont des changements qui restituent l'allèle de type sauvage.

Des mutations peuvent conduire à une perte de fonction d'un gène ou à une nouvelle fonction.

Les mutations dans des cellules de la lignée germinale peuvent être transmises aux descendants, ce qui n'est pas le cas des mutations somatiques.

On peut disséquer génétiquement un processus biologique si on est en mesure d'obtenir des mutations qui affectent ce processus.



Des colonies mutantes de la moisissure *Aspergillus*. Les mutations impliquées se trouvent dans des gènes qui contrôlent le type et la quantité de plusieurs pigments différents synthétisés par ce champignon, dont la couleur normale est le vert foncé.

(J. Peberdy, Department of Life Sciences, Université de Nottingham, England.)

L'analyse génétique serait impossible sans l'existence de *variants* — des organismes qui présentent des différences phénotypiques pour un caractère donné. Dans les chapitres précédents, nous avons étudié de nombreuses analyses de la transmission de tels variants. Nous allons maintenant en considérer l'origine. Comment les variants génétiques apparaissent-ils ?

La réponse simple à cette question est que les organismes ont une tendance intrinsèque à passer d'un état héréditaire à un autre. Un tel changement transmissible s'appelle une **mutation**. Comme nous le verrons, le terme de mutation couvre une vaste gamme de changements. Dans le chapitre suivant, nous nous intéresserons aux changements mutationnels qui affectent des chromosomes entiers ou de grandes parties de ceux-ci. Dans ce chapitre, nous verrons les événements mutationnels qui se produisent à l'intérieur de gènes individuels. On appelle ce type d'événements, des **mutations géniques**.

Ce chapitre traite de deux sujets essentiels. Premièrement, nous allons envisager les différentes sortes de mutations géniques qui peuvent se produire au niveau de l'ADN et les conséquences de ces changements de l'ADN sur la structure et la fonction des protéines. Les mécanismes moléculaires par lesquels ces mutations apparaissent seront des éléments essentiels de cette discussion. La capacité de certains agents à augmenter le taux de mutation est particulièrement intéressante. Ces agents sont utiles expérimentalement, mais l'intérêt qu'on leur porte vient aussi de la peur de l'homme d'être exposé à de tels agents dans sa vie quotidienne. Il nous faut également considérer les mécanismes moléculaires possibles pour les mutations « spontanées », bien moins fréquentes mais beaucoup plus importantes sur le plan de l'évolution. Dans cette discussion, nous verrons que les cellules ont élaboré des systèmes sophistiqués de réparation de l'ADN, pour empêcher l'apparition de mutations. On peut donc considérer l'ADN comme soumis à une lutte acharnée entre les processus qui peuvent conduire à de nouvelles mutations dans l'ADN et les processus cellulaires de réparation qui les recherchent constamment et éliminent les lésions (ou événements pré-mutationnels). Le deuxième sujet important de ce chapitre est la méthodologie de l'analyse des mutations. Chaque mutation obtenue par l'expérimentateur représente un composant du processus biologique étudié. Nous verrons comment une mutation spécifique intéressante est prélevée parmi de nombreux autres changements dus à des mutations. Ce qui est essentiel dans ce processus est de pouvoir reconnaître les phénotypes aberrants qui révéleront les mutations les plus intéressantes pour l'étude en cours.

LES FONDEMENTS MOLÉCULAIRES DE LA MUTATION

Une mutation est le changement d'un allèle d'un gène en un allèle différent. Puisqu'un tel changement se produit à l'intérieur d'un gène unique et est situé au niveau d'un locus (en un « point ») du chromosome, une mutation dans un gène est parfois appelée **mutation ponctuelle**. Cette

terminologie a été créée avant l'avènement du séquençage de l'ADN et donc avant qu'il soit possible de découvrir de façon aisée la cause moléculaire d'un événement mutationnel. De nos jours, les mutations ponctuelles font référence à des modifications d'une paire de bases ou de quelques paires de bases adjacentes dans l'ADN. Dans ce chapitre, nous nous intéresserons à ce type de mutations ponctuelles simples.

Le nombre de façons dont des mutations ponctuelles peuvent changer un gène de type sauvage est colossal et varie selon la structure et la séquence d'un gène. Cependant, il est toujours vrai que les mutations qui réduisent ou éliminent la fonction d'un gène (mutations dites « perte-de-fonction ») constituent la classe la plus abondante. La raison en est simple : il est bien plus facile de casser une machine que de modifier la façon dont elle fonctionne en changeant ou en enlevant au hasard un de ses composants. Pour la même raison, les mutations qui augmentent ou modifient le type d'activité du gène ou l'endroit où celui-ci est exprimé (mutations dites « gain-de-fonction ») sont bien plus rares.

Les changements au niveau de l'ADN

Les mutations ponctuelles sont classifiées en termes moléculaires dans le tableau 7-1. Ce dernier présente les principaux types de changements dans l'ADN et leurs conséquences fonctionnelles au niveau protéique.

Il existe deux grands types de changements mutationnels ponctuels dans l'ADN : les *substitutions de bases* et les *additions ou délétions de bases*. Les **substitutions de bases** sont les mutations dans lesquelles une paire de bases est remplacée par une autre. Les substitutions de bases peuvent à leur tour être divisées en deux sous-classes : les transitions et les transversions. Pour décrire ces sous-classes, nous allons considérer la façon dont une mutation modifie la séquence d'un brin d'ADN (le changement complémentaire aura lieu dans l'autre brin). Une **transition** est un remplacement d'une base par l'autre base de la même catégorie chimique (une purine remplacée par une purine : soit A en G ou G en A ; une pyrimidine remplacée par une pyrimidine : soit C en T ou T en C). Une **transversion** est l'opposé — le remplacement d'une base appartenant à une catégorie chimique par une base appartenant à l'autre catégorie (une pyrimidine remplacée par une purine : C en A, C en G, T en A, T en G ; une purine remplacée par une pyrimidine : A en C, A en T, G en C, G en T). Lorsqu'on décrit les mêmes changements pour les ADN double-brin, il faut spécifier les deux membres d'une paire de bases : un exemple de transition serait : G•C → A•T ; celui d'une transversion serait G•C → T•A.

Les **mutations par addition ou délétion** concernent en réalité des paires de *nucléotides* ; quoi qu'il en soit, par convention, on les appelle des additions ou des délétions de paires de *bases*. Les plus simples de ces mutations sont des additions ou des délétions d'une seule paire de bases. Il existe des exemples dans lesquels des mutations apparaissent à la suite d'additions ou de délétions

Hidden page

Hidden page

l'on appelle une **substitution synonyme**, alors il est probable que la modification aura un effet moins sévère sur la structure et la fonction de la protéine. À l'inverse, le remplacement d'un acide aminé par un autre chimiquement différent, ce que l'on appelle une **substitution non synonyme**, est davantage susceptible de produire des changements importants dans la structure et la fonction de la protéine concernée. Les mutations non-sens conduisent à la terminaison prématurée de la traduction. Elles ont donc un effet considérable sur la fonction de la protéine. Les mutations non-sens produisent en général des protéines complètement inactives, sauf si elles se produisent très près de l'extrémité 3' du cadre de lecture ouvert et donnent lieu à un polypeptide tronqué en partie fonctionnel.

Comme les mutations non-sens, les additions ou délétions d'un seule base ont des conséquences sur la séquence polypeptidique, qui s'étendent bien au-delà du site de mutation lui-même. La séquence d'ARNm étant « lue » par la machinerie de traduction par groupes de trois paires de bases (codons), l'addition ou la délétion d'une seule paire de bases dans l'ADN modifiera le cadre de lecture à partir de la position de cette addition ou délétion et cela, jusqu'à l'extrémité carboxy-terminale de la protéine. On appelle donc ces lésions, des mutations par **décalage du cadre de lecture**. À la suite de ce type de mutations, la totalité de la séquence d'acides aminés traduits, situés en aval du site mutant, n'a rien à voir avec la séquence d'acides aminés originale. Les mutations par décalage du cadre de lecture entraînent généralement la perte intégrale de la structure et de la fonction de la protéine concernée.

Tournons-nous à présent vers les mutations qui se produisent dans les séquences régulatrices et autres séquences non codantes. Les parties d'un gène qui ne codent pas de protéine contiennent différents sites fonctionnels essentiels. Au niveau de l'ADN, il existe des sites auxquels des protéines spécifiques régulatrices de la transcription doivent se lier. Au niveau de l'ARN, il y a également des séquences fonctionnelles importantes telles que les sites de liaison des ribosomes sur les ARNm bactériens et les sites de ligature utilisés après l'excision des introns dans les ARNm eucaryotes.

Les conséquences des mutations dans les parties d'un gène qui ne codent pas de polypeptide sont bien plus difficiles à prédire. En général, les conséquences fonctionnelles de toute mutation ponctuelle (substitution, addition ou délétion) dans ce type de région dépendent de sa position et du fait qu'elle perturbe ou non un site fonctionnel. Les mutations qui perturbent ces sites sont capables de changer le mode d'expression d'un gène, par exemple la quantité de produit exprimée à un certain moment, ou en réponse à des facteurs environnementaux ou encore dans certains tissus. Nous verrons de nombreux autres exemples de ce type de sites cibles lorsque nous étudierons les mécanismes de la régulation des gènes (chapitres 14-17). Il est important de comprendre que ce type de mutations régulatrices modifie la quantité de produit protéique d'un gène, mais ne change pas la structure de la protéine. Par ailleurs, certaines mutations peuvent complètement inactiver la fonction (comme la fixation de la polymérase

ou l'excision des introns) et être létales.

Manifestement, les gènes contiennent également des séquences non codantes qui ne peuvent subir de mutations ponctuelles produisant des phénotypes détectables. Ces séquences sont entrecoupées de sites mutables. Elles peuvent ne jouer aucun rôle fonctionnel ou bien d'une façon ou d'une autre, être protégées des lésions causées par les mutations.

Les nouvelles mutations sont répertoriées comme *induites* ou *spontanées*. Les **mutations induites** sont celles qui apparaissent à la suite d'un traitement délibéré par des mutagènes, c'est-à-dire des agents de l'environnement qui augmentent le taux de mutations. Les **mutations spontanées** apparaissent en l'absence de traitement mutagène connu. Elles constituent le « bruit de fond » mutationnel et sont sans doute la première cause de variation génétique naturelle que l'on observe dans les populations.

La fréquence à laquelle les mutations spontanées apparaissent est faible, généralement d'une cellule sur 10^5 à 10^6 . Si un grand nombre de mutants sont nécessaires pour une analyse génétique, les mutations doivent être *induites*. L'induction des mutations se fait en traitant les cellules par des mutagènes. Les mutagènes les plus couramment utilisés sont les radiations à haute énergie ou certains produits chimiques spécifiques. Des exemples de ces mutagènes et de leur efficacité sont donnés dans le tableau 7-2, à la page suivante. Plus la dose de mutagène est élevée, plus le nombre de mutations induites est important, comme on le voit dans la figure 7-1. Remarquez

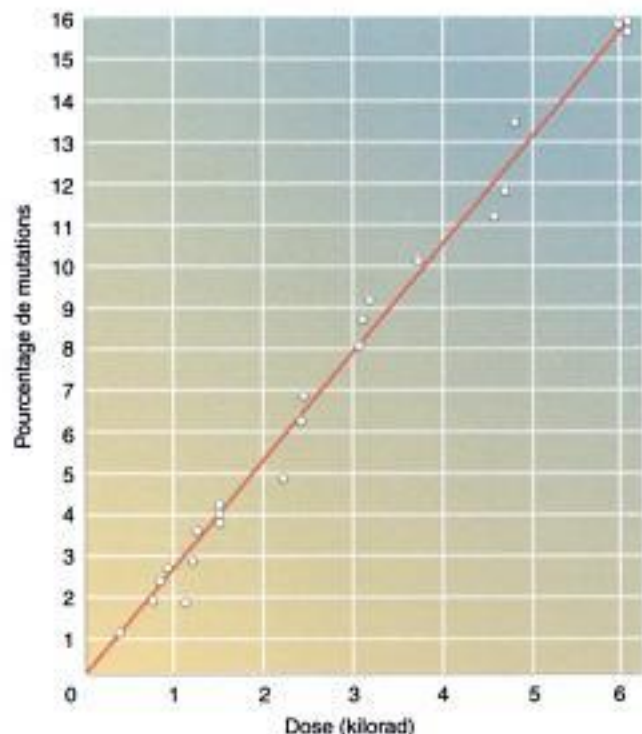


Figure 7-1 Relation linéaire entre la dose de rayons X à laquelle des mouches *Drosophila melanogaster* ont été exposées et le pourcentage de mutations (essentiellement des mutations létales récessives liées au sexe).

TABLEAU 7-2

Fréquences de mutations directes obtenues avec divers mutagènes chez *Neurospora*

Traitement mutagène	Temps d'exposition (minutes)	Survie (%)	Nombre de mutants <i>ad-3</i> pour 10 ⁶ survivants
Pas de traitement (taux spontané)	—	100	~ 0,4
Amino-purine (1-5 mg/ml)	Pendant la croissance	100	3
Éthyl méthane sulfonate (1 %)	90	56	25
Acide nitreux (0,05 M)	160	23	128
Rayons X (2000 r/mn)	18	16	259
Méthyl méthane sulfonate (20 mM)	300	26	350
Rayonnements UV (600 erg/mm ² par mn)	6	18	375
Nitrosoguanidine (25 µM)	240	65	1500
Moutarde d'acridine ICR-170 (5 µg/ml)	480	28	2287

NOTE : Le test mesure la fréquence de mutants *ad-3*. Il s'avère que ces mutants sont rouges, ce qui permet de les distinguer sur le fond blanc des colonies *ad-3*⁺.

que la figure 7-1 présente une réponse *linéaire* par rapport à la dose de mutagène, ce que l'on observe souvent lors de l'induction de mutations ponctuelles. Les mécanismes moléculaires par lesquels les mutagènes agissent seront traités dans les paragraphes suivants.

La distinction entre mutation induite et spontanée dépend en réalité de la procédure expérimentale utilisée. Si nous savons qu'un organisme a été **soumis à un agent mutagène**, alors nous déduisons que toute mutation apparue après cette mutagenèse a été induite. Cependant cela n'est pas tout à fait vrai. Les mécanismes qui provoquent des mutations spontanées sont toujours en action dans un organisme qui a subi un traitement mutagène. En réalité, une partie des mutations observées après la mutagenèse seront indépendantes de celle-ci. La proportion de mutations qui s'inscrivent dans cette catégorie dépend de la puissance de l'agent mutagène. Plus le taux de mutations induites est élevé, plus la proportion de mutations observées qui sont effectivement « spontanées » est faible.

Les mutations induites et spontanées apparaissent généralement à la suite de mécanismes différents. Nous les examinerons séparément. Après avoir considéré ces mécanismes, nous envisagerons la réparation biologique des mutations. Sans ces mécanismes de réparation, le taux de mutation serait si élevé que les cellules accumuleraient trop de mutations pour être viables et pour pouvoir se

reproduire. Les événements mutationnels qui se produisent sont donc les rares événements qui, d'une façon ou d'une autre, ont été oubliés ou contournés par les processus de réparation.

Les mécanismes d'induction des mutations

Lorsque nous examinons la gamme des mutations induites par des mutagènes différents, nous pouvons voir une spécificité distincte, caractéristique de chaque mutagène. Une telle **spécificité mutationnelle** a été remarquée pour la première fois au niveau du locus *rII* du bactériophage T4. Cette spécificité vient d'une préférence d'un mutagène donné, à la fois pour un *type* de mutation (par exemple, des transitions G•C→A•T) et pour certains *sites* mutationnels appelés **points chauds**.

Les mutagènes agissent par le biais d'au moins trois mécanismes différents. Ils peuvent *remplacer* une base de l'ADN, *modifier* une base de façon à entraîner un mésappariement spécifique avec une autre base, ou *endommager* une base qui ne peut plus s'apparier avec aucune base dans des conditions normales.

Le remplacement d'une base. Certains composés chimiques ressemblent suffisamment aux bases azotées

Hidden page

Hidden page

avec la cytosine dans les réplifications ultérieures. Ou bien, si la 2-AP est incorporée par un mésappariement avec la cytosine, il y aura des transitions $G \cdot C \rightarrow A \cdot T$ lorsqu'elle s'appariera avec la thymine. Des études génétiques ont montré que la 2-AP, comme le 5-BU, est hautement spécifique des transitions.

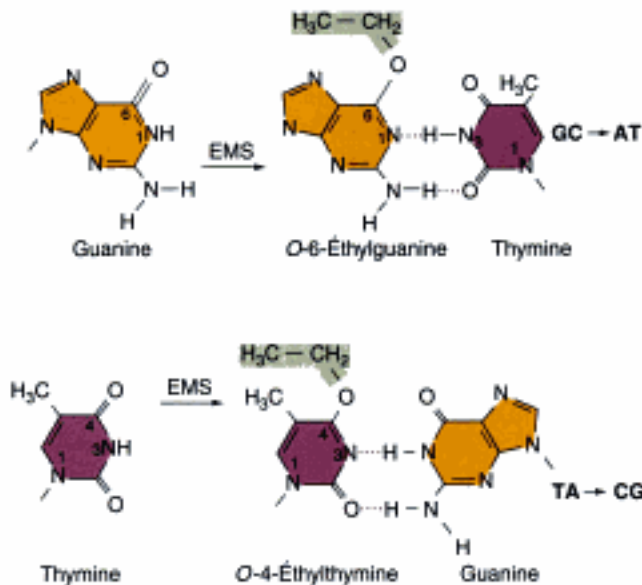
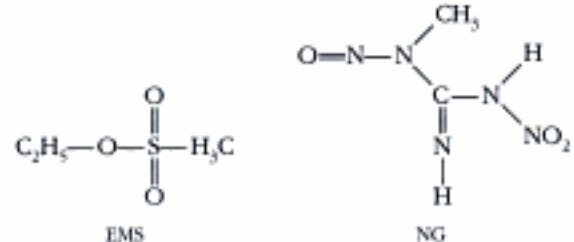


Figure 7-7 Mésappariement spécifique induit par l'alkylation. L'alkylation (dans ce cas, l'éthylation due à l'EMS) de la position O-6 de la guanine, ainsi que de la position O-4 de la thymine, peut conduire à un mésappariement respectivement avec la thymine et la guanine, comme on le voit ici. Chez les bactéries, où les mutations ont été analysées plus en détail, les principales mutations détectées sont des transitions $G \cdot C \rightarrow A \cdot T$, indiquant que l'alkylation en O-6 de la guanine est plus sensible à la mutagenèse.

Les modifications des bases. Certains mutagènes ne sont pas incorporés dans l'ADN et au lieu de cela modifient une base, en provoquant un mésappariement spécifique. Certains **agents alkylants**, tel que l'**éthyl méthane-sulfonate (EMS)** et la **nitrosoguanidine (NG)** largement utilisée, fonctionnent de cette façon :



Bien que de tels agents ajoutent des groupements alkyle (un groupement éthyle dans le cas de l'EMS et un groupement méthyle dans le cas de la NG) en de nombreuses positions des quatre bases, la mutagenicité conduit le plus souvent à une addition sur l'oxygène en position 6 de la guanine pour créer une O-6-alkylguanine. Cette alkylation conduit au mésappariement direct avec la thymine, comme le montre la figure 7-7 et aboutit à des transitions $G \cdot C \rightarrow A \cdot T$ au cycle suivant de réplification. Les agents alkylants peuvent également modifier les bases des nucléotides entrants en cours de synthèse d'ADN.

Les **agents intercalants** constituent une autre classe de modificateurs d'ADN. Ce groupe de composés comprend la **proflavine**, l'**acridine orange** et une classe de composés chimiques appelés **composés ICR** (figure 7-8a). Ces agents sont des molécules planes qui ressemblent aux paires de bases et sont capables de se glisser (de *s'intercaler*) entre les bases azotées empilées, au cœur de la double hélice d'ADN (figure 7-8b). Dans cette position, un agent

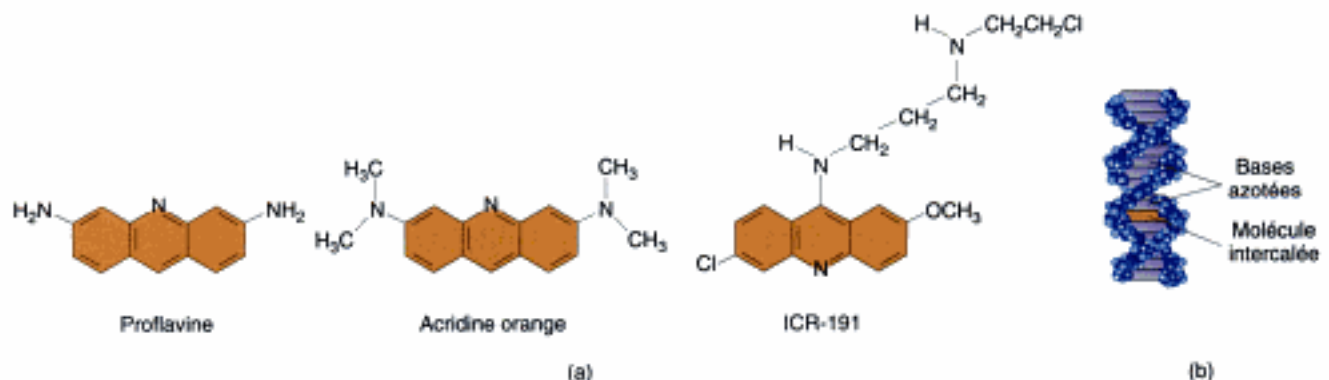


Figure 7-8 Agents intercalants. (a) Structures des agents usuels proflavine, acridine orange et ICR-191. (b) Un agent intercalant se glisse entre les bases azotées empilées dans la molécule d'ADN. Cet événement peut conduire à des insertions et des délétions d'une seule paire de nucléotides. (D'après L. S. Lerman, *Proceedings of the National Academy of Sciences USA* 49, 1963, 94).

peut provoquer des insertions ou des délétions d'une seule paire de nucléotides. Les agents intercalants peuvent également s'empiler entre les bases, dans l'ADN simple-brin. De cette façon, ils peuvent stabiliser les bases, qui ont formé des boucles lors du décalage du cadre de lecture, comme le décrit le modèle de Streisinger (figure 7-9).

Les lésions des bases. Un grand nombre de mutagènes endommagent une ou plusieurs bases, empêchant ensuite tout appariement spécifique de bases. Le résultat est un blocage de la réplication, car la synthèse de l'ADN ne peut se poursuivre au-delà d'une base qui ne peut reconnaître* son partenaire complémentaire par la formation de liaisons hydrogène. Dans les cellules bactériennes, ce type d'obstacle à la réplication peut être *franchi* en insérant des bases non spécifiques. Le processus nécessite l'activation d'un système spécial, le **système SOS** (figure 7-10). On appelle ce système *SOS* car il apparaît comme une solution de secours pour empêcher la mort de la cellule en présence d'une lésion importante de l'ADN. L'induction SOS se fait en dernier ressort, ce qui permet à la cellule d'échapper à la mort au prix d'un certain niveau de mutagenèse.

On ne connaît pas exactement le fonctionnement du mécanisme de franchissement (*by-pass*) SOS, bien que l'on sache que chez *E. coli*, il dépend d'au moins trois gènes, *recA* (qui joue également un rôle dans la recombinaison), *umuC* et *umuD*. Les modèles actuels de ce système SOS suggèrent que les protéines UmuC et UmuD se combinent avec le complexe de réplication d'ADN polymérase III, pour relâcher sa spécificité normalement très stricte et permettre à la réplication de se poursuivre au-delà des séquences lésées.

La figure 7-10 présente un modèle du mécanisme de franchissement qui intervient après l'arrêt de la polymérase III au niveau d'un type de lésion appelé photodimère T-C. Puisque la réplication peut recommencer après le dimère, une région simple-brin d'ADN est produite. Ceci attire une protéine stabilisatrice appelée protéine de liaison aux simples-brins (SSB), ainsi que la protéine RecA, qui forme des filaments autour de l'ADN simple-brin et déclenche la synthèse des protéines UmuC et UmuD. La protéine UmuD se fixe aux filaments et est clivée par la protéine RecA, ce qui produit une version tronquée appelée UmuD', qui recrute ensuite la protéine UmuC avec laquelle elle forme un complexe. Celui-ci permet à l'ADN polymérase III de continuer son chemin au-delà du dimère, en introduisant en face de celui-ci des bases avec une fréquence d'erreur élevée (voir Figure 7-10).

L'action des mutagènes qui endommagent des sites spécifiques d'appariement de bases dépend donc du système SOS. La catégorie de mutagènes SOS-dépendants est importante car elle contient la plupart des agents provoquant les cancers (carcinogènes), tels que la lumière ultraviolette (UV) et l'aflatoxine B₁. Comment le système SOS participe-t-il à la réparation des mutations après la mutagenèse ? Le système SOS abaisse-t-il à tel point la fidélité de la réplication d'ADN (pour permettre le franchissement des lésions) que de nombreuses erreurs de réplication se produisent, même dans l'ADN non endommagé ? Si cette hypothèse était correcte, la plupart des mutations produites par les différents mutagènes SOS-dépendants seraient similaires, plutôt que spécifiques de chacun. La majorité des mutations résulteraient de l'action du système SOS lui-même sur l'ADN non endommagé. Le mutagène

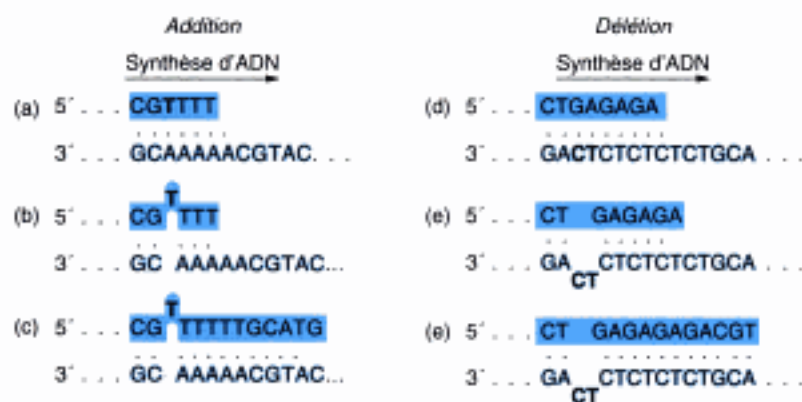


Figure 7-9 Une version simplifiée du modèle de Streisinger pour la formation d'un décalage du cadre de lecture. (a) à (c) Durant la synthèse d'ADN, le brin néo-synthétisé glisse, formant à l'extérieur du brin une boucle composée d'une ou plusieurs bases. Cette boucle est stabilisée par l'appariement permis par l'unité de séquence répétitive (les bases A dans ce cas). Il y aura addition d'une paire de bases, A•T, au cycle suivant de réplication dans cet exemple. (d) à (f) Si, au lieu du brin néo-synthétisé, c'est le brin matrice qui glisse, alors une délétion se produira. Ici, l'unité répétée est un dinucléotide C•T. Après le glissement, une délétion de deux paires de bases (C•G et T•A) aura lieu au cycle de réplication suivant.

* Note du réviseur : Certaines modifications de base ne bloquent pas la réplication mais la polymérase provoque une erreur en incorporant une base erronée.

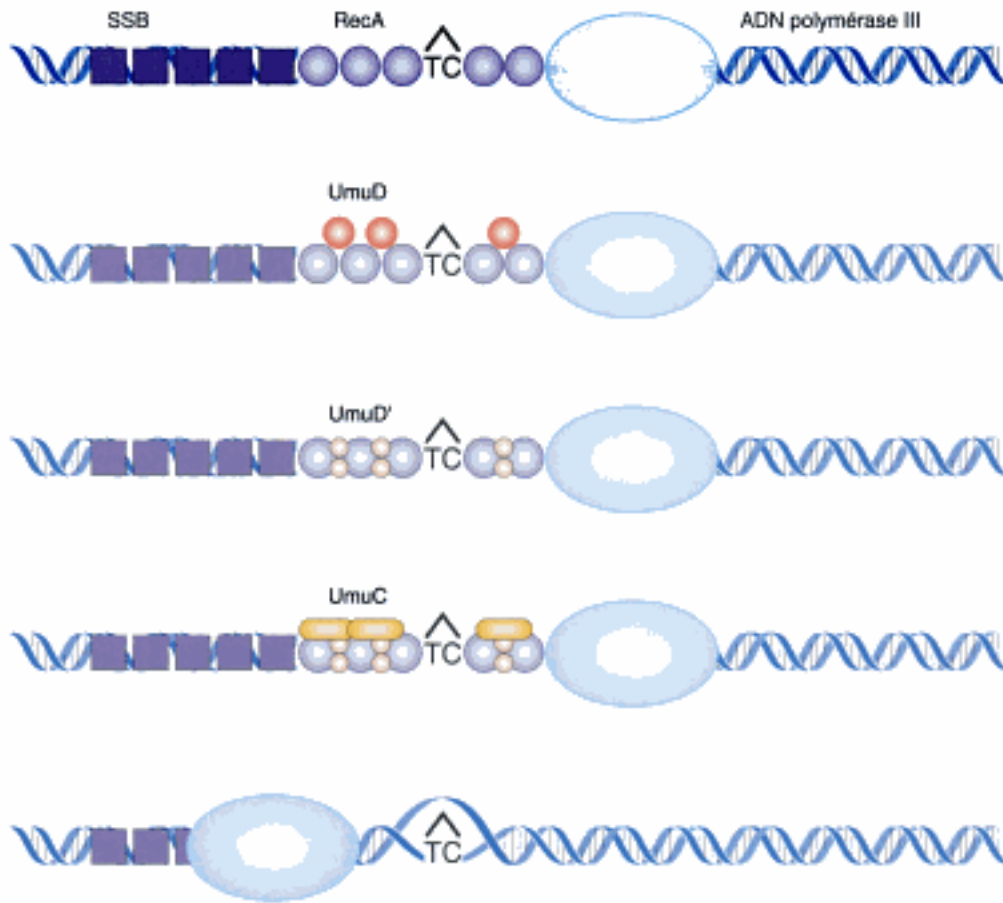


Figure 7-10 L'ADN polymérase, en bleu, s'arrête au niveau d'une lésion non codante, telle que le photodimère T•C présenté ici, formant des régions simple-brin qui attirent les protéines SSB (violet) et RecA (mauve), qui forme des filaments. La présence des filaments de RecA aide à signaler à la cellule la nécessité de synthétiser UmuD (cercles rouges), qui est clivée par RecA en UmuD' (cercles roses) et UmuC (ovales jaunes). UmuC est recrutée pour former un complexe avec UmuD', qui permet à l'ADN polymérase III de poursuivre son action au-delà de la lésion bloquante.

jouerait alors un rôle indirect dans l'induction du système SOS. Les études de spécificité des mutations ont montré que ce n'est pas le cas. Au lieu de cela, différents mutagènes SOS-dépendants présentent des spécificités distinctes. Chaque mutagène induit une distribution unique de mutations. Les mutations doivent donc être produites en réponse à des paires spécifiques de bases endommagées. Le type de lésion diffère dans de nombreux cas. Les photoproduits dus aux UV et les sites apuriniques sont parmi les lésions les plus étudiées.

La **lumière ultraviolette (UV)** induit de nombreux photoproduits dans l'ADN. Deux lésions différentes qui relient des pyrimidines adjacentes dans le même brin sont fortement corrélées avec la mutagenèse. Ces lésions sont

le photodimère cyclobutane pyrimidine et le photoproduit 6-4 (figure 7-11 à la page suivante). Ces lésions interfèrent avec l'appariement normal des bases. L'induction du système SOS est donc nécessaire à la mutagenèse. L'insertion de bases incorrectes en face de photoproduits créés par les UV se fait en position 3' du dimère, et plus fréquemment, pour des dimères 5'-CC-3' et 5'-TC-3'. La transition C → T est la mutation la plus fréquente, mais d'autres substitutions de bases (transversions) et décalages de cadres de lecture sont également induits par la lumière ultraviolette, ainsi que des délétions et duplications plus importantes.

L'**aflatoxine B₁ (AFB₁)** est un puissant carcinogène, isolé pour la première fois à partir de cacahuètes infectées par des champignons. L'aflatoxine est ajoutée en position

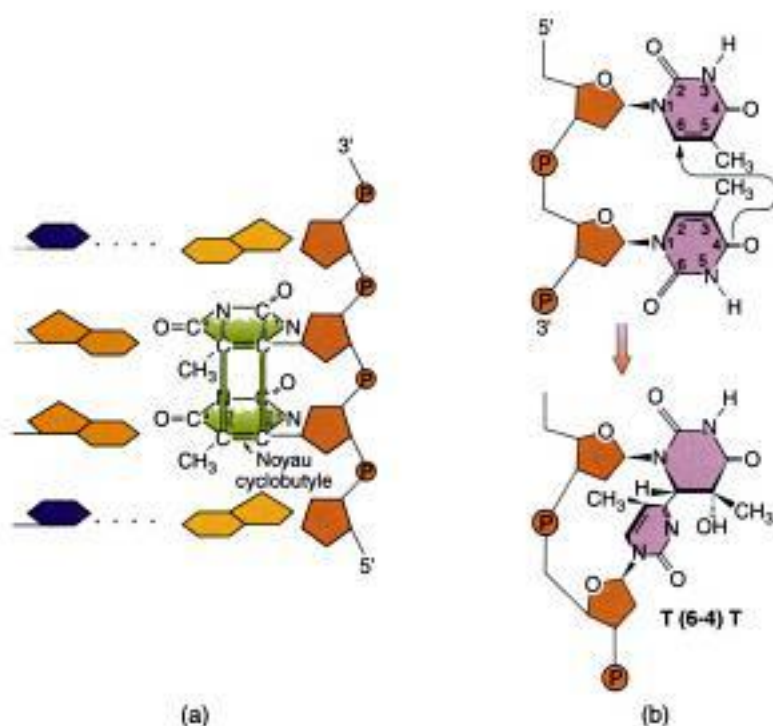


Figure 7-11 (a) Structure d'un dimère de cyclobutane pyrimidine. La lumière ultraviolette stimule la formation d'un noyau cyclobutyle à quatre bras (en vert) entre deux pyrimidines adjacentes sur le même brin d'ADN, en agissant sur les double liaisons 5,6. (b) Structure du photoproduit 6-4. La structure se forme de préférence avec 5'-C-C-3' et 5'-T-C-3', entre les positions C-6 et C-4 de deux pyrimidines adjacentes, entraînant une perturbation importante de la structure locale de la double hélice. (Partie a adaptée de E. C. Friedberg, *DNA repair*. Copyright © 1985 par W. H. Freeman. Partie b d'après J. S. Taylor et coll.)

N-7 de la guanine (figure 7-12). Ce produit conduit à la rupture de la liaison entre la base et le sucre, libérant ainsi la base et formant un **site apurinique** (figure 7-13). Les études menées sur les sites apuriques produits *in vitro* ont démontré que le franchissement de ces sites grâce au système SOS conduisait à l'insertion préférentielle

d'une adénine en face d'un site apurinique. Ceci laisse prévoir que les agents responsables de dépurination au niveau de résidus guanine devraient induire préférentiellement des transversions $G \cdot C \rightarrow T \cdot A$. Par exemple, avec le 0 représentant un site apurinique :

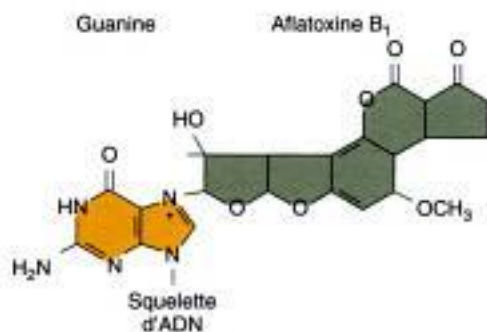
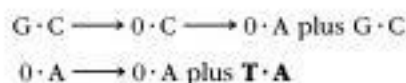


Figure 7-12 La liaison à l'ADN de l'aflatoxine B₁ activée par le métabolisme.

MESSAGE

Les mutagènes induisent des mutations par différents mécanismes. Certains mutagènes ressemblent aux bases normales et sont incorporés dans l'ADN, où ils peuvent former des mésappariements. D'autres endommagent des bases et entraînent un mésappariement ou détruisent un appariement existant en supprimant la reconnaissance entre les bases.

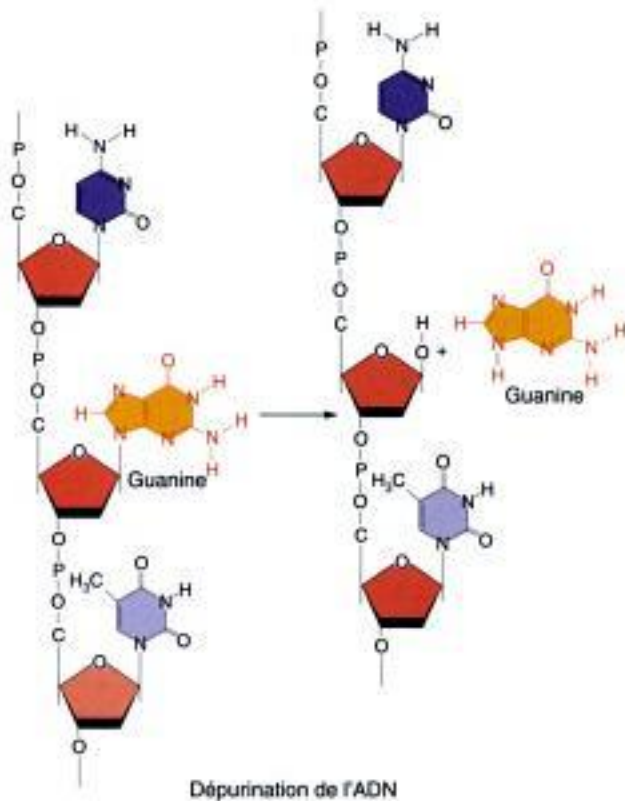


Figure 7-13 La perte d'un résidu purine (guanine) d'un simple-brin d'ADN. Le squelette sucre-phosphate reste intact.

Les mécanismes de mutation spontanée

L'origine des changements héréditaires spontanés a toujours suscité un vif intérêt (voir Les grandes découvertes de la génétique 7-1). On sait maintenant que les mutations spontanées proviennent de sources très diverses, y compris d'erreurs dans la réplication de l'ADN, de lésions spontanées et d'éléments génétiques transposables. Les deux premières sources sont traitées dans ce paragraphe, la troisième sera examinée au chapitre 13.

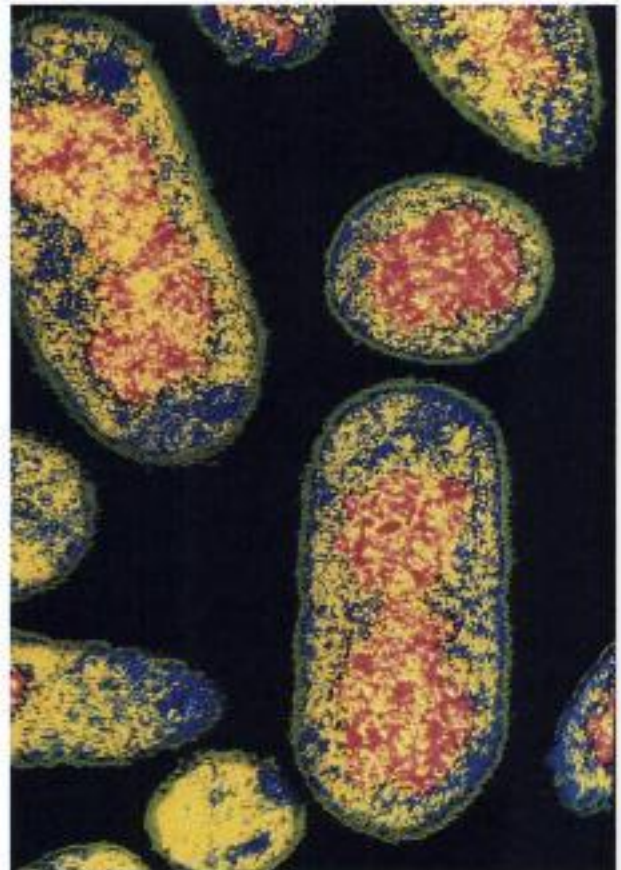
Les mutations spontanées sont très rares, ce qui rend difficile l'examen des mécanismes sous-jacents à leur apparition. Comment pouvons-nous donc avoir une idée des processus gouvernant la mutation spontanée ? Même si ces mutations sont rares, certains systèmes sélectifs permettent d'obtenir de nombreuses mutations spontanées et de les caractériser ensuite au niveau moléculaire — par exemple, en déterminant leur séquence d'ADN. D'après la nature des changements de séquence, on peut déduire certaines informations à propos des processus responsables de ces mutations spontanées.

LES GRANDES DÉCOUVERTES DE LA GÉNÉTIQUE 7-1

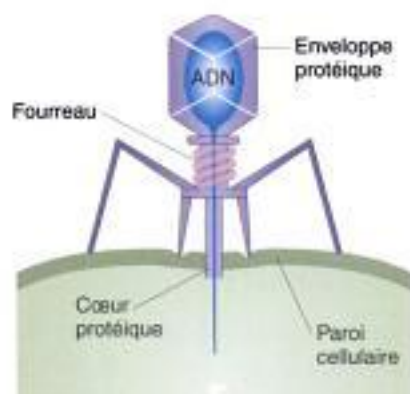
Salvador Luria et Max Delbrück montrent que les mutations chez les bactéries sont aléatoires

Les bactéries et les phages ont joué des rôles essentiels dans la recherche génétique, y compris dans l'étude des mutations. Une expérience réalisée en 1943 par Luria et Delbrück, sur des bactéries et des phages, a particulièrement influencé notre compréhension de la nature de la mutation, non seulement chez les bactéries, mais chez les organismes en général.

On savait à cette époque que si l'on étalait des bactéries *Escherichia coli* sur une boîte contenant du milieu nutritif, en présence d'un phage appelé T1, les bactéries étaient rapidement parasitées et tuées par les phages. Toutefois, rarement mais régulièrement, on observait certaines colonies résistantes à l'attaque des phages. Ces colonies étaient stables et apparaissaient donc comme de véritables mutants. Cependant, on ne savait pas si ces mutants étaient produits spontanément mais à n'importe quel moment ou s'ils étaient induits par la présence du phage.



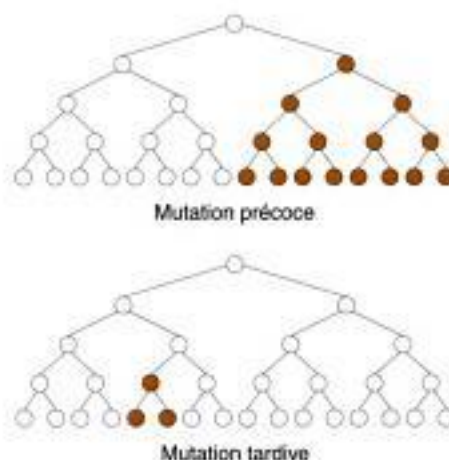
A. B. Dowsett/Science Photo Library/Photo Researchers.



Luria déduisit que si les mutations se produisaient spontanément, alors on devait s'attendre à observer des mutations à des moments distincts dans des cultures différentes, de sorte que le nombre résultant de colonies résistantes par culture devrait présenter une variation élevée (ou « fluctuation » selon ses propres termes). Il affirma plus tard que l'idée lui était venue en regardant les gains variables obtenus par des collègues qui jouaient aux machines à sous dans un club de loisirs local.



Luria et Delbruck construisirent leur « test de fluctuation » de la façon suivante. Ils inoculèrent vingt petites cultures, constituées de quelques cellules chacune et les incubèrent jusqu'à obtenir une concentration de 10^8 cellules par millilitre. En parallèle, une culture bien plus importante était inoculée et incubée jusqu'à parvenir à la même concentration. Les 20 cultures individuelles et 20 aliquots de la même taille issus de la culture de grande taille initiale furent étalés sur une boîte de culture en présence du phage. Les 20 cultures individuelles présentèrent une variation élevée du nombre de colonies résistantes : 11 boîtes avaient 0 colonie résistante et les autres en avaient 1, 1, 3, 5, 5, 6, 35, 64 et 107. Les 20 aliquots issus de la culture importante présentaient bien moins de variation d'une boîte à l'autre.

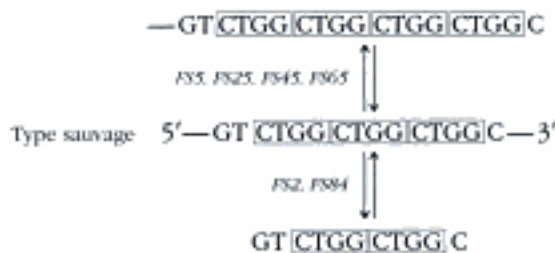


le nombre de colonies résistantes étant dans tous les cas compris entre 14 et 26. Si le phage induisait les mutations, il n'y aurait aucune raison pour que la fluctuation soit plus élevée dans les cultures individuelles, car toutes étaient exposées au phage de la même manière. La meilleure interprétation était que la mutation s'était produite aléatoirement dans le temps : les mutations précoces donnaient des nombres plus élevés de cellules résistantes car elles avaient le temps de produire de nombreux descendants résistants. Les dernières mutations produisaient moins de cellules résistantes. Ce résultat conduisit au « paradigme » actuel de la mutation, c'est-à-dire que chez les virus, les bactéries ou les Eucaryotes, une mutation peut apparaître dans n'importe quelle cellule, à n'importe quel moment.

Les données du test de fluctuation permirent également de calculer le taux de mutation par division cellulaire et cette approche est toujours utilisée en génétique microbienne.

Les erreurs au cours de la réplication de l'ADN. La formation de mésappariements au cours de la réplication est une source de substitution spontanée de bases. (Nous avons déjà parlé des mésappariements au cours de la discussion sur le 5-BU). La plupart des mutations causées par des mésappariements sont des transitions. Ceci est sans doute dû au fait qu'un mésappariement A•C ou G•T ne déforme pas autant la double hélice qu'un appariement A•G ou C•T. Cependant, des transversions peuvent également apparaître à la suite d'un mésappariement. Les erreurs de réplication peuvent également conduire à des mutations par décalage du cadre de lecture. La séquence nucléotidique entourant les points chauds où sont concentrées les mutations par décalage du cadre de lecture a été déterminée dans le gène du lysozyme du phage T4. Ces mutations ont souvent lieu au niveau de bases répétées. Selon le modèle de Streisinger (figure 7-9), le décalage du cadre de lecture apparaît lorsque des boucles formées dans des régions simple-brin sont stabilisées par

l'appariement décalé des séquences répétées au cours de la réplication. De plus, dans le gène *lacI* d'*E. coli*, certains points chauds apparaissent au niveau de séquences répétées, exactement comme le prédit le modèle de Streisinger. La figure 7-14 représente la distribution des mutations spontanées dans le gène *lacI*. Remarquez de quelle façon un ou deux sites de mutations dominent la distribution. Dans *lacI*, une séquence de quatre paires de bases (CTGG), répétée trois fois en tandem dans le type sauvage, est la cause des points chauds (pour plus de simplicité, seul un brin d'ADN est représenté) :



Le principal point chaud, représenté ici par les mutations FS5, FS25, FS45 et FS65 résulte de l'addition d'un groupe supplémentaire des quatre bases CTGG. Le point chaud secondaire, représenté ici par les mutations FS2 et FS84 résulte de la perte d'un groupe de ces quatre bases CTGG.

Comment pouvons-nous expliquer ces observations ? Le modèle de Streisinger prédit que la fréquence d'un décalage du cadre de lecture donné dépend du nombre de paires de bases qui peuvent se former au cours de l'appariement décalé des séquences répétées. La séquence de type sauvage présentée pour le gène *lacI* peut glisser et exposer une séquence CTGG et stabiliser cette structure en formant neuf paires de bases (appliquez le modèle de la figure 7-9 à la séquence présentée pour *lacI*). Si le glissement a lieu respectivement sur le brin matrice ou sur le brin néo-synthétisé, il s'ensuivra une délétion ou une addition.

Les délétions plus importantes (de plus de quelques bases) constituent une fraction assez élevée des mutations spontanées, comme le montre la figure 7-14. La plupart

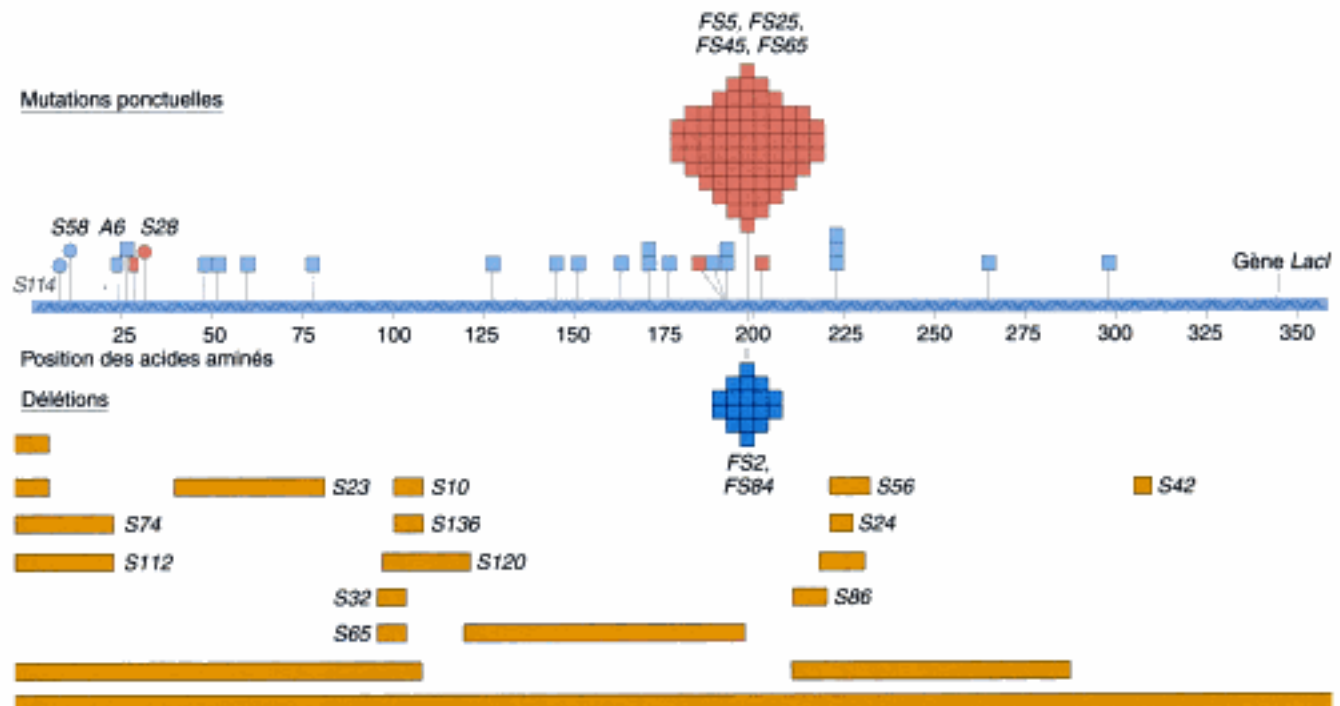


Figure 7-14 Distribution de 140 mutations spontanées dans *lacI*. Chaque événement correspondant à une mutation ponctuelle est indiqué par un carré. Les carrés orange correspondent aux mutations qui réversent facilement. Les délétions (en jaune) sont représentées dessous. Les nombres de la carte du gène *I* correspondent à la position des acides aminés du répresseur Lac codé par *I*. Les numéros d'allèles font référence aux mutations qui ont été analysées au niveau de la séquence d'ADN. Les mutations S114 et S58 (cercles bleus) résultent de l'insertion d'éléments transposables (voir chapitre 13). S28 (cercle orange) est une duplication de 88 paires de bases. (D'après P. J. Farabaugh, U. Schmeissner, M. Hofer et J. H. Miller, *Journal of Molecular Biology* 126, 1978, 847.)

des délétions (mais pas toutes) concernent la perte de séquences répétées. La figure 7-15 montre les 12 premières délétions analysées au niveau de la séquence d'ADN. D'autres études ont montré que les répétitions plus longues constituent des points chauds de délétions. On a observé des duplications de segments d'ADN chez de nombreux organismes. Comme les délétions, elle se produisent souvent au niveau de répétitions de séquences.

Comment ces délétions et ces duplications apparaissent-elles ? Plusieurs mécanismes pourraient expliquer leur formation. Les délétions pourraient être le résultat d'erreurs de réplication. Par exemple, une extension du modèle de Streisinger pour l'appariement décalé (figure 7-14) pourrait expliquer pourquoi les délétions prédominent au niveau de courtes séquences répétées. Par ailleurs, les délétions et duplications pourraient être produites par recombinaison entre des répétitions.

Les lésions spontanées. Les **lésions spontanées**, qui sont des lésions naturelles de l'ADN, peuvent également créer des mutations. Deux des lésions spontanées les plus courantes sont la désamination et la dépurination, cette dernière étant la plus fréquente.

Nous avons vu précédemment que l'aflatoxine induisait la **dépurination**. Toutefois, la dépurination se produit également spontanément. Une cellule de mammifère perd spontanément 10 000 purines de son ADN pendant une période de division de 20 heures à 37°C. Si ces lésions persistaient, elles entraîneraient des dégâts génétiques importants car, au cours de la réplication, les sites apuriques ne peuvent spécifier aucun type de base, *a fortiori* pas la base correcte. Cependant, comme nous l'avons

S74, S112 75 bases			
CAATTCAGG GTGGTGA ATGTGAAACC-----CGC GTGGTGA ACCAGG			
Site (N° des pb)	Répétition de séquence	Nbre de bases délétées	Nombre d'événements (et leur appellation)
20 à 95	GTGGTGA	75	2 S74, S112
146 à 269	GCGGCGAT	123	1 S23
331 à 351	AAGCGGCG	20	2 S10, S136
316 à 338	GTCGA	22	2 S32, S65
694 à 707	CA	13	1 S24
694 à 719	CA	25	1 S56
943 à 956	G	13	1 S42
322 à 393	Aucune	71	1 S120
658 à 685	Aucune	27	1 S86

Figure 7-15 Les délétions dans *lacI*. Les délétions qui se produisent dans S74 et S112 sont indiquées en haut du tableau. Comme le montrent les barres jaunes, l'une des répétitions de séquence (en vert) et tout l'ADN intermédiaire ont été délétés, laissant une copie de la séquence répétée. Toutes les mutations ont été analysées par détermination directe de la séquence d'ADN. (D'après P. J. Farabaugh, U. Schmeissner, M. Hofer et J. H. Miller, *Journal of Molecular Biology* 126, 1978, 847.)

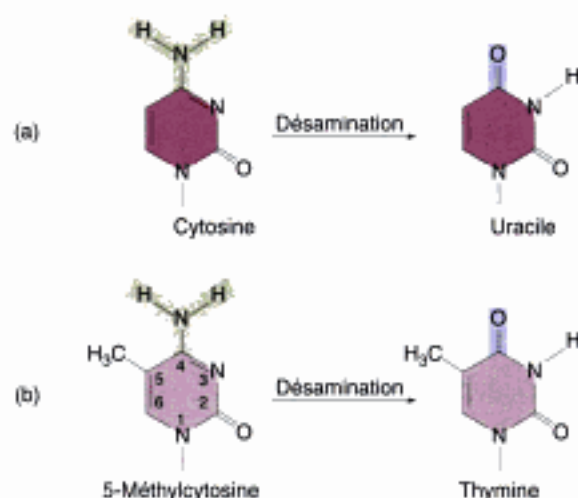


Figure 7-16 Désamination de (a) la cytosine et (b) de la 5-méthylcytosine.

mentionné plus haut dans ce chapitre, dans certaines conditions, une base peut être insérée en face d'un site apurinique, ce qui aboutit fréquemment à une mutation.

La **désamination** de la cytosine produit l'uracile (figure 7-16a). Les résidus uracile non réparés s'apparieront avec l'adénine pendant la réplication, ce qui aboutira à la conversion d'une paire G•C en une paire A•T (une transition G•C → A•T). On a découvert qu'à certains locus, la désamination des cytosines expliquait un type de point chaud de mutations. L'analyse de la séquence d'ADN des points chauds pour les transitions G•C → A•T dans le gène *lacI* a montré que les résidus 5-méthylcytosine sont présents au niveau de chaque point chaud. (Certaines bases chez les Procaryotes et les Eucaryotes sont normalement méthylées). Certaines des données issues de cette étude de *lacI* sont indiquées dans la figure 7-17. La hauteur de chaque barre sur le graphique représente la fréquence de mutations au niveau de chacun des nombreux sites. On peut voir que les positions des résidus 5-méthylcytosine sont bien corrélées avec la plupart des sites mutables.

Comment des 5-méthylcytosines peuvent-elles conduire à des mutations ? L'une des enzymes de réparation présentes dans la cellule, l'uracile-ADN glycosylase, reconnaît dans l'ADN les résidus uracile qui proviennent de désaminations et les excise, laissant une brèche qui est ensuite comblée. Cependant, la désamination de la 5-méthylcytosine (figure 7-16b) produit de la thymine (5-méthyluracile) qui n'est pas reconnue par l'uracile-ADN glycosylase et n'est donc pas réparée. On observe donc plus fréquemment des transitions C → T produites par désamination au niveau des sites 5-méthylcytosine, car elles échappent à ce système de réparation.

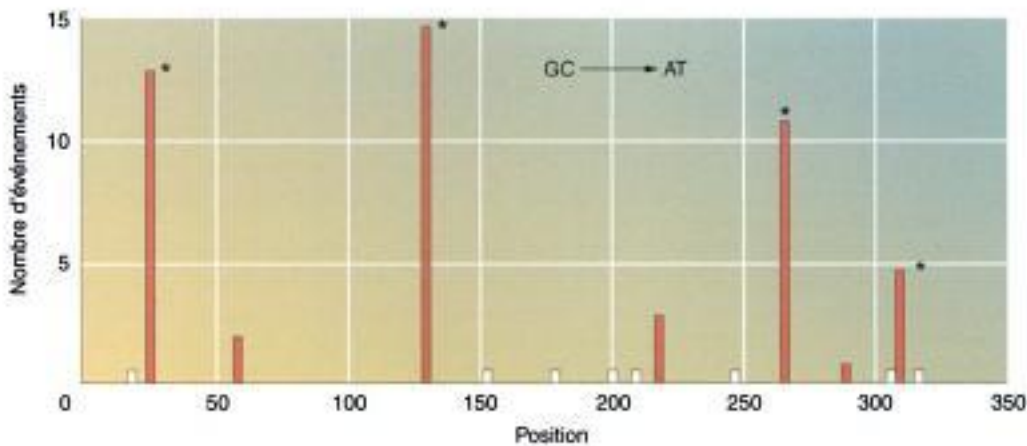


Figure 7-17 Les points chauds de la 5-méthylcytosine chez *E. coli*. Les mutations non-sens qui se produisent en 15 sites différents de *lacI* ont été notées. Toutes résultent d'une transition G•C → A•T. Les astérisques (*) signalent les positions des 5-méthylcytosines. Les colonnes blanches représentent des sites au niveau desquels un changement G•C → A•T pourrait être détecté mais où aucune mutation de ce type n'a eu lieu. On peut voir que les résidus 5-méthylcytosine constituent des points chauds pour la transition G•C → A•T. Sur 50 mutations apparues indépendamment, 44 étaient situées au niveau des quatre sites de 5-méthylcytosine et seulement 6 au niveau des 11 cytosines non méthylées. (D'après C. Coulondre et coll., *Nature* 274, 1978, 775).

Les **bases endommagées par oxydation** constituent un troisième type de lésion spontanée conduisant à une mutation. Les formes d'oxygène actif, telles que les radicaux superoxyde ($O_2\bullet$), le peroxyde d'hydrogène (H_2O_2) et les radicaux hydroxyle ($OH\bullet$) sont des sous-produits du métabolisme aérobie normal. Ces molécules peuvent provoquer des dégâts oxydatifs dans l'ADN, ainsi que dans les précurseurs de l'ADN (tels que le GTP), ce qui crée des mutations. Ce type de mutations a été mis en cause dans de nombreuses maladies génétiques humaines. La figure 7-18 montre deux produits de lésion oxydative. La 8-oxo-7-hydrodésoxyguanosine (8-oxodG, ou •GO•) forme fréquemment un mésappariement avec A, ce qui aboutit à un niveau élevé de transversions G → T.

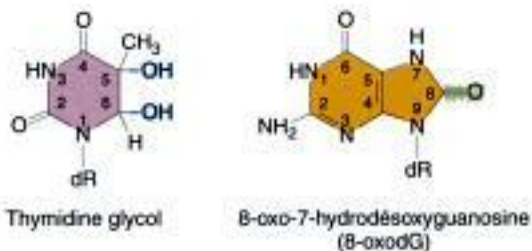


Figure 7-18 Bases endommagées à la suite de l'attaque de l'ADN par des radicaux oxygène. dR = désoxyribose.

MESSAGE

Les mutations spontanées peuvent être produites par plusieurs processus différents. Les erreurs de réplication et les lésions spontanées créent la plupart des mutations par substitution de base ou décalage du cadre de lecture.

Les mutations spontanées et les maladies humaines.

L'analyse de la séquence d'ADN a révélé des lésions mutationnelles responsables d'un grand nombre de maladies humaines héréditaires. Beaucoup sont les substitutions de bases, addition ou délétion simples attendues. Certaines sont plus complexes mais rappellent les mutations bactériennes dont nous avons parlé précédemment, ce qui nous permet de suggérer des mécanismes susceptibles de provoquer ces maladies humaines.

Un grand nombre de ces affections sont dues à des délétions ou des duplications de séquences répétées. Par exemple, les encéphalomyopathies mitochondriales sont un groupe de maladies affectant le système nerveux central ou les muscles (par exemple, le syndrome de Kearns-Sayre). Elles se caractérisent par un dysfonctionnement de la phosphorylation oxydative mitochondriale et par des changements dans la structure de l'ADN mitochondrial. On a montré que ces affections résultaient de délétions de séquences d'ADN situées entre des séquences répétées. La figure 7-19 représente l'une de ces délétions. Remarquez

Hidden page

Hidden page

Dans quelques cas cependant, les lésions peuvent être réparées de cette façon. Un photodimère mutagène résultant de l'action de la lumière UV en est un exemple (figure 7-11). Le photodimère de cyclobutane pyrimidine peut être réparé par une **photolyase** que l'on a trouvée chez des bactéries et quelques Eucaryotes inférieurs, mais pas chez l'homme. L'enzyme se fixe au photodimère et le scinde en deux en présence de certaines longueurs d'onde de la lumière visible, pour redonner les bases d'origine (figure 7-21). Cette enzyme ne peut agir dans le noir, aussi, d'autres voies de réparation sont-elles nécessaires pour supprimer ces dégâts dus aux UV. On a également détecté une photolyase qui inverse les photoproduits 6-4 chez des végétaux et chez la drosophile.

Les **alkyltransférases** sont également des enzymes qui inversent les lésions. Elles enlèvent les groupements alkyle qui ont été ajoutés en position O-6 de la guanine (figure 7-7) par des mutagènes tels que la nitrosoguanidine et l'éthyl méthanesulfonate. La méthyltransférase d'*E. coli* a été étudiée en détail. Cette enzyme transfère le groupement méthyle de l'O-6-méthylguanine à un résidu cystéine de la protéine. Lorsque cela se produit, l'enzyme est inactivée, de sorte que ce système de réparation peut être saturé si le niveau d'alkylation est suffisamment élevé.

Les systèmes de réparation par excision. Le système général de **réparation par excision** casse une liaison phosphodiester de part et d'autre de la lésion, sur le même brin, entraînant ainsi l'excision d'un oligonucléotide.

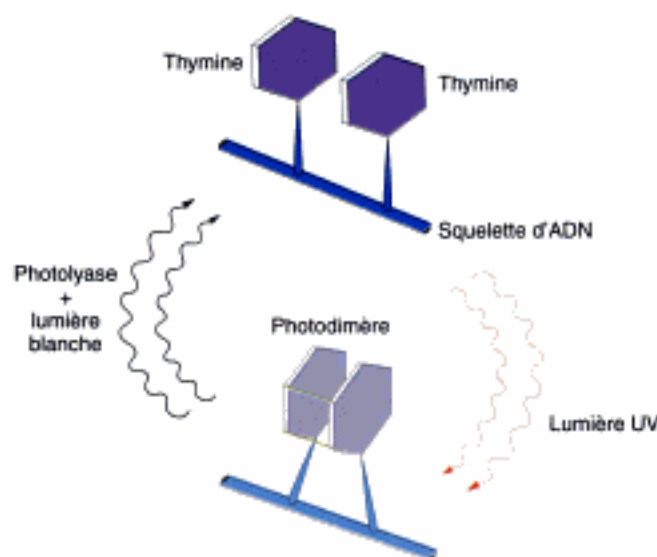


Figure 7-21 La réparation d'un photodimère de pyrimidine induit par des UV, par une enzyme photoréactive ou photolyase. L'enzyme reconnaît le photodimère (ici un dimère de thymine) et s'y fixe. La photolyase utilise l'énergie de la lumière, lorsqu'il y en a, pour scinder le photodimère en deux, redonnant ainsi les monomères d'origine. (D'après J. D. Watson, *Molecular biology of the gene*, 3^e éd. Copyright © 1976 par W. A. Benjamin).

Ceci laisse une brèche, remplie par une synthèse réparatrice ; une ligase ressoude ensuite les cassures. Chez les Procaryotes, 12 ou 13 nucléotides sont retirés, alors que chez les Eucaryotes, 27 à 29 nucléotides sont éliminés. La figure 7-22 présente le déroulement de l'excision dans chacun des cas.

Certaines lésions sont trop fines pour entraîner une déformation suffisamment importante qui leur permettrait d'être reconnues par le système général de réparation par excision et par son équivalent dans les cellules supérieures. D'autres **systèmes spécifiques d'excision** sont donc nécessaires. La **réparation par excision de bases** est réalisée par des **ADN glycosylases** qui clivent les liaisons N-glycosidiques (base-sucre), libérant ainsi les bases endommagées et produisant des sites apuriques ou apyrimidiques (sites AP ; voir figure 7-13). La première étape de ce processus est présentée dans la figure 7-23. Le site résultant est ensuite réparé par un système de réparation endonucléasique, spécifique des sites AP.

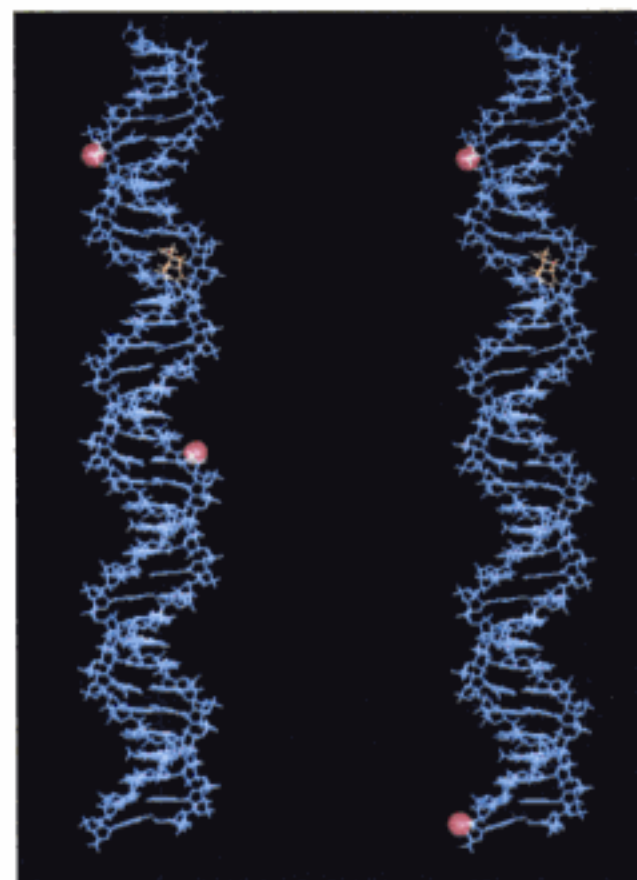


Figure 7-22 Exemples d'incision excinuéasique par des enzymes d'*E. coli* (à gauche) et des enzymes humaines. Les points rouges indiquent les endroits incisés pour une lésion, dans ce cas un dimère de thymine, figuré en orange. (Merci à J. E. Hearst dans A. Sancar, *Science* 266, 1974, 1954).

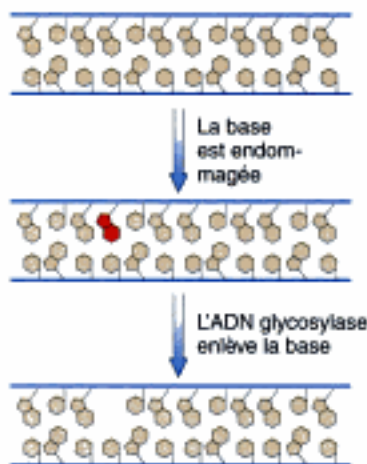


Figure 7-23 Action des ADN glycosylases. La glycosylase retire une base endommagée et laisse un site AP. Le site AP est ensuite excisé par les endonucléases AP schématisées dans la figure 7-24. (D'après B. Lewin, *Genes*. Copyright © 1983 par John Wiley).

Il existe de nombreuses ADN glycosylases. L'une d'entre elles, l'uracil-ADN glycosylase retire l'uracile de l'ADN. Les résidus uracile, qui apparaissent à la suite de la désamination spontanée de la cytosine (figure 7-16), peuvent conduire à une transition $C \rightarrow T$ s'ils ne sont pas réparés. Il est possible que le partenaire d'appariement normal de l'adénine dans l'ADN soit la thymine (5-méthyluracile) plutôt que l'uracile, pour permettre la reconnaissance et l'excision de ces résidus uracile. Si l'uracile était un composant normal de l'ADN, alors une telle réparation serait impossible.

Toutes les cellules possèdent des endonucléases qui attaquent les sites créés par la perte spontanée d'un seul résidu de purine ou de pyrimidine. Les **endonucléases AP** sont vitales pour la cellule car, comme nous l'avons vu précédemment, la dépurination spontanée est un événement relativement fréquent. Ces enzymes pratiquent des coupures dans les chaînes en clivant les liaisons phosphodiester au niveau des sites AP. Ceci initie un processus de réparation par excision qui met en jeu trois autres enzymes — une exonucléase, une ADN polymérase I et une ADN ligase (figure 7-24).

Étant donné l'efficacité de la voie de réparation des endonucléases AP, elle peut servir d'étape finale à d'autres voies de réparation. Si des paires de bases endommagées peuvent être excisées, laissant un site AP, les endonucléases AP peuvent alors terminer la restauration du type sauvage. C'est ce qui se passe dans la voie de réparation des ADN glycosylases.

Réparation post-répliationnelle. Certains systèmes de réparation sont capables de reconnaître des erreurs, même après la réplication de l'ADN. Un exemple, appelé

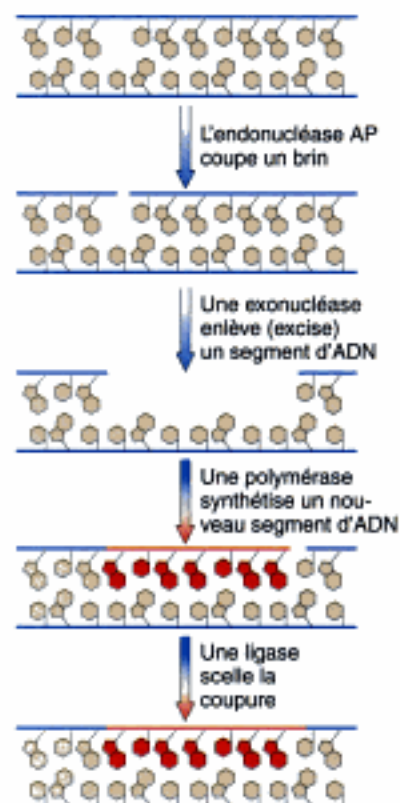


Figure 7-24 Réparation de sites AP (apuriniques ou apyrimidiques). Les endonucléases AP reconnaissent les sites AP et coupent la liaison phosphodiester. Un segment d'ADN est enlevé par une exonucléase et la brèche résultante est comblée par l'ADN polymérase I et refermée par l'ADN ligase. (D'après B. Lewin, *Genes*. Copyright © 1983 par John Wiley).

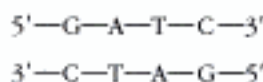
système de **réparation des mésappariements**, peut détecter ce type d'erreurs. Ces systèmes ont au moins trois tâches à accomplir :

1. Reconnaître les paires de bases mal appariées.
2. Déterminer la base incorrecte à l'intérieur du mésappariement.
3. Exciser la base incorrecte et effectuer une synthèse de réparation.

La deuxième propriété de ce système est essentielle. À moins d'être capable de distinguer les bases correctes des bases incorrectes, le système de réparation des mésappariements ne peut déterminer la base à exciser pour empêcher l'apparition d'une mutation. Si par exemple, un mésappariement G•T se produit à la suite d'une erreur de réplication, comment le système peut-il déterminer qui de G ou de T est la base incorrecte ?

Toutes deux sont des bases normales de l'ADN. Toutefois, les erreurs de réplication produisent des mésappariements sur le brin néosynthétisé. C'est donc la base de ce brin, qui doit être reconnue et excisée.

Pour différencier le brin ancien (matrice) du brin néosynthétisé, le système de réparation des mésappariements, mieux connu chez les bactéries, se sert du retard dans la méthylation de la séquence suivante, méthylation qui a lieu normalement après la réplication :



L'enzyme responsable de la méthylation est l'**adénine méthylase**, qui crée une 6-méthyladénine sur chaque brin. Il faut cependant à l'adénine méthylase plusieurs minutes pour reconnaître et modifier les séquences GATC néo-synthétisées. Durant cet intervalle, le système de réparation des mésappariements peut fonctionner, puisqu'il peut à ce moment distinguer le vieux brin du nouveau grâce à la présence ou à l'absence de méthylation. La méthylation en position 6 de l'adénine n'affecte pas l'appariement et fournit un marquage commode, qui peut être détecté par d'autres systèmes enzymatiques. La figure 7-25 montre la fourche de réplication au cours de la correction d'un mésappariement. Remarquez que juste après la réplication, seul le brin ancien est méthylé au niveau des séquences GATC. Une fois le mésappariement identifié, le système de réparation corrige l'erreur.

Le gène *recA* d'*E. coli*, l'un des gènes du système SOS (figure 7-10) participe également à la réparation post-réplicationnelle. Dans ce cas, le système de réparation de l'ADN s'arrête au niveau d'un photodimère causé par les UV, ou d'autres lésions bloquantes, puis reprend sa progression après le blocage, laissant une brèche simple-brin. Le produit RecA participe à la **réparation par recombinaison**, un processus au cours duquel la brèche est comblée par un ADN coupé dans la molécule sœur (figure 7-26a). Ce processus n'est pas sujet à erreur.

Les stratégies de la réparation. Nous pouvons maintenant nous faire une idée de la stratégie des systèmes de réparation utilisés par la cellule. Les nombreux systèmes différents de réparation, disponibles dans la cellule sont résumés dans le tableau 7-3, à la page 220. Ce serait simple si les enzymes pouvaient être utilisées pour inverser directement chaque lésion spécifique. Pourtant c'est parfois impossible pour des raisons chimiques et toutes les lésions possibles de l'ADN ne peuvent être anticipées. Un système général de réparation par excision est donc utilisé pour enlever n'importe quel type de base endommagée qui entraîne une déformation identifiable de la double hélice. Lorsque les lésions sont trop fines pour provoquer une telle déformation, des systèmes spécifiques d'excision, tels que les glycosylases ou des sys-

tèmes de retrait, sont conçus. Pour éliminer les erreurs de réplication, un système de réparation post-réplicationnel des mésappariements intervient. Pour finir, les systèmes de recombinaison post-réplicationnels éliminent les brèches situées en face des lésions bloquantes qui ont échappé aux autres systèmes de réparation.

MESSAGE

Les enzymes de réparation jouent un rôle crucial dans la diminution du nombre de lésions génétiques dans les cellules vivantes. Les cellules possèdent de nombreuses voies de réparation différentes sous leur contrôle pour éliminer des erreurs potentiellement mutagènes.

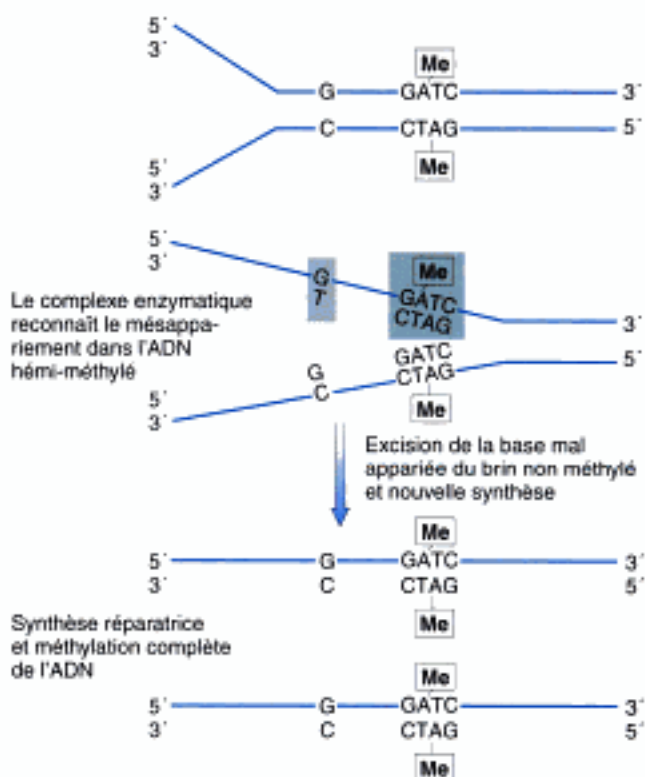
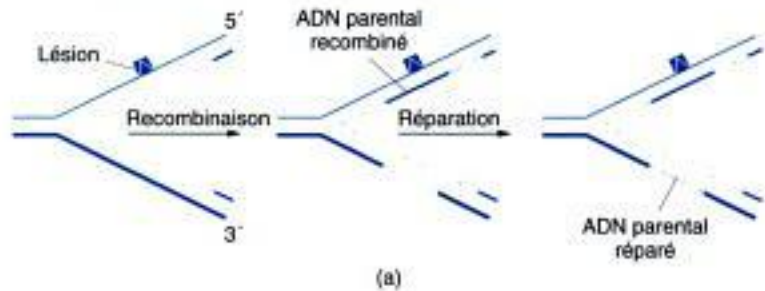


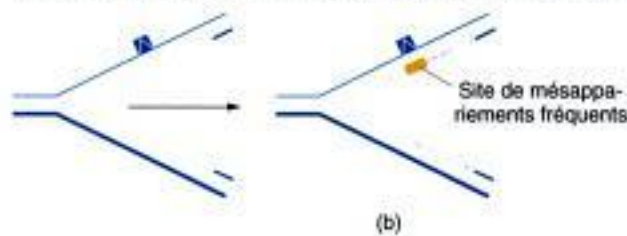
Figure 7-25 Un modèle de réparation des mésappariements chez *E. coli*. L'ADN étant méthylé par des réactions enzymatiques qui reconnaissent le A dans les séquences GATC, juste après la réplication de l'ADN, le brin néo-synthétisé ne sera pas méthylé. L'ADN double-brin « hémiméthylé » permet au système de réparation des mésappariements de distinguer le brin ancien du nouveau brin. Ici, un mésappariement G•T est présenté. Le système de réparation des mésappariements peut reconnaître cette erreur et s'y fixer, déterminer le brin correct (ancien) car c'est lui qui est méthylé dans le double-brin hémiméthylé, puis exciser la base mal appariée du nouveau brin. La synthèse réparatrice restaure les paires de bases normales. (Adapté de E. C. Friedberg, *DNA repair*. Copyright © 1985 par W. H. Freeman and Company, New York).

Figure 7-26 Schémas de la réparation post-répliationnelle. (a) Dans la réparation par recombinaison, la réplication dépasse une lésion bloquante, mais en laissant une brèche dans le nouveau brin. Une protéine dirigée par *recA* comble ensuite cette brèche, en utilisant un morceau du brin parental opposé (en raison de la complémentarité de l'ADN, ce morceau fournira les bases correctes pour la brèche). Pour finir, la protéine RecA répare la brèche dans le brin parental. (b) Dans le système de franchissement SOS, lorsque la réplication atteint une lésion bloquante, le système SOS insère le nombre de bases nécessaire (souvent des bases incorrectes) en face de la lésion et la réplication continue sans laisser de brèche. Remarquez que dans les deux cas, la lésion d'origine est toujours là et doit être réparée par un autre système de réparation. (Adapté de A. Kornberg et T. Baker, *DNA replication*, 2^e édition. Copyright © 1992 par W. H. Freeman and Company, New York).

Réparation post-répliationnelle par recombinaison



Réplication prédisposée aux erreurs (franchissement de lésion) : Système SOS



Les erreurs de réparation et les maladies humaines.

On sait ou on suppose que plusieurs maladies génétiques humaines sont causées par la déficience des gènes intervenant dans les systèmes de réparation. Ces défauts augmentent souvent la fréquence d'apparition des cancers. En effet, en raison de l'augmentation générale du taux de mutation, les gènes sont plus fréquemment mutés, ce qui peut amener une cellule à la cancérisation (voir chapitre 15).

L'une des maladies qui augmente les risques de can-

cer, appelée *xeroderma pigmentosum* (XP), résulte d'un défaut dans n'importe lequel des huit gènes impliqués dans la réparation par excision de nucléotides. Les personnes atteintes de cette maladie sont extrêmement sujettes aux cancers de la peau induits par les UV (figure 7-27) à la suite de l'exposition au soleil et présentent des anomalies neurologiques fréquentes. La différence de photosensibilité aux UV entre les cellules normales et les cellules XP apparaît clairement lorsqu'on regarde les courbes de survie de la figure 7-28.



Figure 7-27 Cancer de la peau dans la maladie *xeroderma pigmentosum*. Cette maladie héréditaire récessive est due à une déficience de l'une des enzymes de réparation par excision, ce qui conduit à la formation de cancers de la peau lors de l'exposition de la peau aux rayonnements UV du soleil. (Photographie aimablement communiquée par Dirk Bootsma, Erasmus University, Rotterdam).

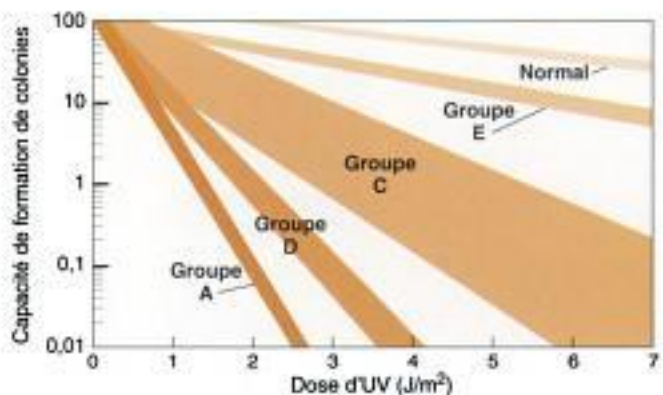


Figure 7-28 Hypersensibilité aux rayonnements UV de cellules XP en culture. Ici sont figurées des cellules issues de plusieurs groupes de complémentation. Il y a une variation entre les différents groupes de complémentation, mais tous sont plus sensibles aux rayonnements UV que les cellules normales (Adapté de J. E. Cleaver et K. H. Kraemer, *Xeroderma pigmentosum*, dans C. R. Scriver et coll., eds. *The metabolic basis of inherited disease*. Copyright © 1989 par McGraw-Hill Book Co., New York).

TABLEAU 7-3
Les systèmes de réparation chez *E. coli*

Mode d'action général	Exemple	Type de lésion réparée	Mécanisme
Détoxification	Superoxyde dismutase	Empêche la formation de lésions oxydatives	Convertit les peroxydes en peroxyde d'hydrogène, neutralisé par la catalase
Élimination directe des lésions	Alkyltransférases	O-6-alkylguanine	Transfère le groupement alkyle de l'O-6-alkylguanine à un résidu cystéine de la transférase
	Photolyase	Photoproduit 6-4	Romp la liaison 6-4 et restaure les bases normales
	Photolyase	Photodimères créés par les UV	Coupe les dimères en présence de lumière blanche
Excision générale	Système exonucléasique codé par <i>uvrABC</i>	Lésions entraînant des déformations de la double hélice, telles que les photoproduits créés par les UV et les additions de groupements chimiques volumineux	Introduit des coupures endonucléolytiques de part et d'autre de la lésion ; la brèche résultante est réparée par l'ADN polymérase I et l'ADN ligase
Excision spécifique	Endonucléases AP	Sites AP	Introduit des coupures endonucléolytiques ; une exonucléase crée une brèche qui est réparée par l'ADN polymérase I et l'ADN ligase
	ADN glycosylases	Bases désaminées (uracile, hypoxanthine), certaines bases méthylées, des purines dont les cycles sont ouverts, des bases ayant subi des lésions oxydatives et certaines autres bases modifiées	Enlève la base, créant un site AP réparé par une endonucléase AP
	Système GO	8-oxodG	Une glycosylase enlève la 8-oxodG de l'ADN ; une autre glycosylase convertit les mésappariements 8-oxodG•A restants en paires 8-oxodG•C et la première glycosylase enlève ensuite la 8-oxodG
Post-réplicationnel	Système de réparation des mésappariements	Erreurs de réplication provoquant des mésappariements de paires de bases	Reconnait le brin néo-synthétisé en y détectant les résidus adénine non méthylés dans les séquences 5'-GATC-3' ; excise ensuite les bases du nouveau brin lorsqu'il détecte un mésappariement
	Réparation par recombinaison	Les lésions qui bloquent la réplication. Laisse des segments simple-brin	Échange par recombinaison
	Système SOS	Lésions qui bloquent la réplication	Permet à la réplication de se poursuivre malgré la lésion bloquante, ce qui provoque souvent l'apparition de mutations en face de la lésion

L'ANALYSE DES MUTATIONS

Comprendre les bases moléculaires des mutations est un préliminaire indispensable à l'analyse par celles-ci, des voies de développement ainsi que des voies biochimiques et de la biologie cellulaire. En général, le chercheur s'intéresse à un processus biologique particulier et souhaite utiliser l'approche génétique pour analyser celui-ci. Le but est d'identifier tous les gènes qui interviennent dans ce processus, puis de comprendre la nature des produits de ces gènes (généralement des protéines) ainsi que la façon dont ils contribuent et interagissent dans ce processus ou cette voie biologique. Les gènes sont identifiés grâce à leurs allèles mutants. L'approche génétique commence donc avec des mutants. Généralement, on suit les étapes indiquées ci-dessous pour disséquer génétiquement un processus :

1. Concevoir un système efficace pour détecter les mutations.
2. Utiliser un mutagène pour induire un grand nombre de mutants qui présentent des variations par rapport au processus de type sauvage.
3. Regrouper les mutations par gènes en utilisant des tests de complémentation.
4. Attribuer aux gènes leur locus chromosomique.
5. Isoler les gènes en utilisant la technologie de l'ADN.
6. Caractériser la structure et la fonction des gènes.
7. À partir de l'analyse de l'interaction des gènes et de l'interaction des protéines, dresser un portrait du fonctionnement du processus biologique étudié.

Nous verrons des exemples détaillés de l'application de ces idées à la génétique du développement, dans les chapitres 15 et 16. Pour le moment, il nous faut considérer les approches générales.

Il y a un gouffre entre les types de mutations que l'on croit *susceptibles* d'apparaître dans un gène et les mutations qu'on récupère vraiment après mutagenèse. Plusieurs facteurs sont essentiels pour obtenir des mutations après une mutagenèse. Le choix du mutagène est important car des mutagènes différents produiront des gammes distinctes de mutations. Les phénotypes utilisés pour identifier les individus porteurs de mutations sont également importants car seule une petite partie de l'ensemble des mutations possibles dans un gène donné conduira à la production du phénotype mutant adéquat. C'est pourquoi, dans toute procédure de mutagenèse, chaque gène représentera une cible d'une taille spécifique. Il est possible qu'un gène constitue une cible relativement grande pour un type de mutagenèse et une cible relativement petite pour un autre type de mutagenèse. Le défi pour le généticien expérimentateur consistera alors à imaginer une mutagenèse qui « saturera » le système, c'est-à-dire, qui permettra d'identifier chaque composant du processus biologique étudié.

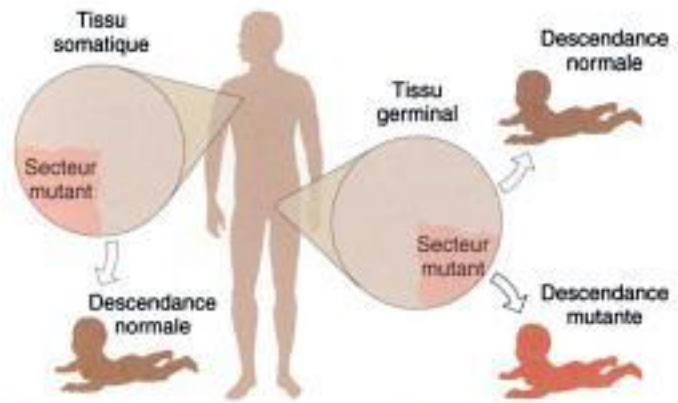


Figure 7-29 Les mutations somatiques ne sont pas transmises à la descendance, tandis que les mutations germinales peuvent être transmises à une partie ou à toute la descendance.

Les mutations somatiques et les mutations germinales

L'une des questions qui se pose lors de la mise au point de l'analyse mutationnelle est le type de tissu ou de cellules à examiner. Chez des organismes pluricellulaires, des gènes peuvent être mutés dans le tissu somatique ou dans le tissu germlinal. On appelle ces changements respectivement des **mutations somatiques** et des **mutations germinales**. Ces deux types sont schématisés dans la figure 7-29. Si une *mutation somatique* se produit dans une seule cellule d'un tissu somatique en cours de développement, cette cellule donnera naissance à une population de cellules mutantes identiques, qui descendront toutes de cette cellule mutée. On appelle **clone** une population de cellules identiques provenant d'une même cellule mère par division asexuée. Les membres d'un clone ayant tendance à rester proches les uns des autres au cours du développement, l'un des résultats observables d'une mutation somatique est souvent une tache de cellules phénotypiquement mutantes, appelée **secteur mutant**. Plus la mutation se produit tôt dans le développement, plus le secteur mutant sera grand (figure 7-30). Les secteurs

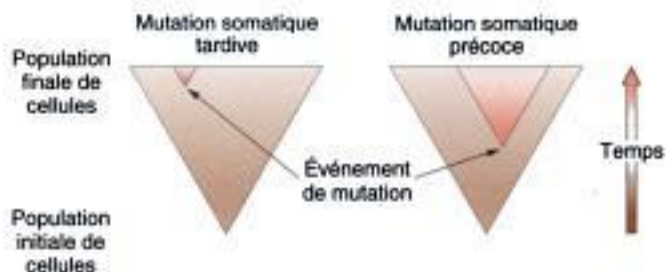


Figure 7-30 Une mutation précoce produit une proportion plus importante de cellules mutantes dans la population en croissance qu'une mutation tardive.

mutants peuvent être identifiés visuellement seulement si leur phénotype contraste suffisamment avec celui des cellules environnantes de type sauvage (figure 7-31).

Chez les diploïdes, on s'attend à ce que toute mutation somatique *dominante* se manifeste dans le phénotype de la cellule ou du clone de cellules qui la contient. À l'inverse, une mutation *récessive* ne sera pas exprimée, car elle sera masquée par un allèle de type sauvage, qui par définition est dominant sur la mutation récessive. Une deuxième mutation pourrait créer une mutation homozygote récessive, mais cet événement est extrêmement rare.

Quelles seraient les conséquences d'une mutation somatique dans une cellule appartenant à un organisme complètement développé ? Si la mutation se produit dans un tissu dans lequel les cellules sont en cours de division, alors il est possible qu'un clone mutant apparaisse. Si la mutation a lieu dans une cellule post-mitotique — c'est-à-dire qui ne se divise plus — alors l'effet sur le phénotype sera très probablement négligeable. Même si des mutations dominantes font mourir une cellule ou la rendent déficiente, cette perte de fonction sera compensée par d'autres cellules normales dans ce tissu. Toutefois, les mutations qui donnent lieu à un cancer sont un cas à part. Les mécanismes du cancer seront traités au chapitre 15.

Les mutations somatiques peuvent-elles être transmises à la descendance ? C'est impossible car les cellules somatiques sont par définition celles qui ne sont jamais transmises à la descendance. Cependant, il est à noter que si l'on prend une bouture de plante, provenant d'une tige ou d'une feuille qui comporte un secteur somatique mutant, la plante qui poussera à partir de la bouture mutante pourra développer du tissu germinale mutant. En d'autres termes, une branche portant des fleurs (c'est-à-dire du tissu germinale) pourra pousser à partir du secteur somatique mutant. Ce qui est apparu comme une mutation somatique pourra donc être transmis de manière sexuée. Un exemple est présenté dans la figure 7-32.



Figure 7-31 Une mutation somatique dans une pomme « red delicious ». L'allèle mutant déterminant la couleur dorée (golden) est apparu dans une cloison ovarienne de la fleur, qui a formé la partie charnue de la pomme. Les graines ne seront pas mutantes et donneront naissance à des pommiers à fruits rouges. (Remarquez qu'en réalité, la variété « golden delicious » a pris naissance sous la forme d'une branche mutante d'un arbre « red delicious »). (Anthony Griffiths).



Figure 7-32 Une mutation produisant un allèle pour des pétales blancs est apparue dans le tissu somatique. Celui-ci étant devenu une partie du tissu germinale, la mutation a pu être transmise par l'intermédiaire des graines. La mutation est apparue dans un bourgeon de branche latérale de rosier. La branche a poussé et a produit des fleurs. (D'après Harper Horticultural Slide Library).

Une *mutation germinale* se produit dans la **lignée germinale**, un tissu particulier qui est isolé au cours du développement et à partir duquel se formeront les gamètes. Si un gamète mutant participe à la fécondation, alors la mutation sera transmise à la génération suivante. Toutefois, un individu avec un phénotype normal et des ancêtres normaux peut avoir des gamètes mutants non détectés. Ces mutations ne peuvent être détectées que si elles se trouvent chez un zygote (figures 7-33 et 7-34). Nous avons vu au chapitre 4 que la mutation de l'hémophilie liée à l'X dans les familles royales d'Europe avait sans doute eu lieu dans les cellules germinales de la reine Victoria ou de ses parents. La mutation n'était exprimée que chez les descendants mâles.

Les systèmes de détection des mutations chez les haploïdes sont souvent très simples. En effet, puisqu'il y a un seul jeu de chromosomes, une mutation est exprimée automatiquement sous l'aspect d'un phénotype mutant. Donc si nous nous intéressons (par exemple) à la voie de synthèse du pigment caroténoïde orange que l'on trouve normalement dans les cellules du champignon *Neurospora*, toute mutation qui inactive cette voie donnera des colonies incolores ou avec un pigment différent qui pourra être un intermédiaire de la voie de synthèse. On distingue facilement de telles colonies à l'œil nu. Examiner un grand nombre d'individus à la recherche d'un mutant s'appelle un **criblage**. Il faut différencier le criblage de la **sélection**, au cours de laquelle le phénotype mutant survit à un traitement sélectif alors que les types sauvages meurent. Par exemple, si nous nous intéressons aux gènes qui codent les filaments d'actine (qui appartiennent au cytosquelette) chez un champignon, nous pouvons sélectionner des mutants en traitant les cellules par la cytochalasine, un inhibiteur qui se lie à l'actine de type sauvage et l'inactive, empêchant ainsi la prolifération cellulaire. Les mutants qui se développent sur la cytochalasine